М.Н. Борисевич, ***заведующий кафедрой***
Учреждение образования «Витебская ордена «Знак почета» государственная академия ветеринарной медицины»
УДК 636.066+636:612.16

Степень корреляционной связи между температурой и пульсом у мелких животных

Вычислен коэффициент корреляции между температурой и пульсом у домашних позвоночных животных (щенков небольших пород) при воспалении перикарда.

The coefficient of the correlation between the temperature and sphygmus of the house hold backboned animals (puppies of small sized breeds) during the inflammation of pericardium has been established.

Одной из важнейших проблем при анализе заболеваний животных является определение взаимосвязи и взаимозависимости признаков. Задача сложная и многопараметричная. По этой причине она малоизучена в ветеринарной медицине. Между тем эти вопросы без труда могут быть решены в рамках регрессионного и корреляционного анализов, хорошо развитых в математике. Оба подхода позволяют не только устанавливать связи между признаками, но и измерять эти связи количественно.

Цель статьи - определение взаимозависимости двух признаков при заболеваниях мелких животных (щенков небольших пород) воспалением перикарда.

Объекты и методы исследования

Эксперимент, данные которого являются предметом дальнейших обсуждений, осуществлялся на базе городской ветеринарной станции. Из числа поступающих сюда животных отбирались щенки небольших пород с характерными признаками воспаления перикарда (всего 250 животных). Для каждого животного регистрировались два параметра: температура тела (°С) и пульс (число ударов в минуту). Отобранные животные сортировались по нескольким интервалам: по температуре было сформировано 7 интервалов (от 37 до 41°С с шагом 0,1) и по пульсу - 8 интервалов (от 80 до 150 с шагом 10). В результате сортировки была построена большая корреляционная решетка (ее фрагмент для 152 животных приведен в таблице 1). Она и служила основой всех последующих вычислений. Из приведенного фрагмента следует, например, что 7 щенков (из общего числа обследованных) имели одновременно температуру тела, принадлежащую интервалу 38,6-38,9°С (среднее значение интервала 38,75°С) и пульс 110-119 ударов в минуту (среднее значение 114,5).

Взаимосвязь названных параметров может быть изучена двумя способами. Первый способ основан на построении линий регрессии. В эксперименте каждому значению x соответствует определенное среднее значение y. Зависимость среднего значения y от данного x можно рассматривать как функцию x. Последняя представляет собой линию регрессии (y по x):

$$\bar{y}_x = f(x). \qquad (1)$$

Точно так же каждому значению соответствует определенное среднее значение x:

$$\bar{x}_y = f(y). \qquad (2)$$

Это вторая линия регрессии (x по y). Линии регрессии имеют определенную доверительную зону. Если по обе стороны от каждого среднего значения x и y отложить их доверительные интервалы, то получим доверительную зону регрессии. Ширина этой зоны и характеризует степень корреляционной связи исследуемых параметров.

Таблица 1. Фрагмент корреляционной решетки, построенной для двух признаков заболевания мелких животных воспалением перикарда

Пульс, *x* (ударов в минуту)	Температура, *y* (°С)					n_i
	38,2-38,5	38,6-38,9	39,0-39,3	39,4-39,7	39,8-40,1	
-	-	-	-	-	-	
100-109	1	8	4	11	13	37
110-119	3	7	6	8	11	35
120-129	2	3	8	21	7	41
130-139	2	4	4	6	4	20
140-149	1	1	2	2	13	19
-	-	-	-	-	-	
n_j	9	23	24	48	48	152

Второй способ оценки взаимосвязи анализируемых параметров основан на вычислении коэффициента корреляции r. Коэффициент корреляции характеризует степень связи между двумя признаками. Его величина колеблется в пределах от 1 до -1. Чем ближе коэффициент корреляции по абсолютному значению (по модулю) к единице, тем теснее линейная связь между рассматриваемыми признаками. Для независимых друг от друга признаков коэффициент корреляции равен нулю. При этом положительное значение коэффициента корреляции указывает на прямую связь между признаками, в этом случае с возрастанием значения одного параметра возрастает значение и другого; при отрицательном значении коэффициента корреляции связь носит противоположный характер – с возрастанием значения одного параметра уменьшается значение и другого.

Для определения коэффициента корреляции вычисляется вначале момент ковариации признаков x и y.

$$cov\{x,y\} = \frac{1}{N}\sum_{i=1}^{n}\sum_{j=1}^{m} n_{ij}(x_i-\bar{x})(y_j-\bar{y}), \quad (3)$$

где $\bar{x}$ и $\bar{y}$ – средние значения параметров для всей выборки объема N. Параметры x и y определяются для одних и тех же объектов. Суммирование распространяется на все клетки корреляционной решетки. Коэффициент корреляции вычисляется по формуле

$$r = \frac{cov\{x,y\}}{\sigma_x \cdot \sigma_y}, \quad (4)$$

где σ_x и σ_y - средние квадратичные отклонения случайных величин x и y. При этом важно различать, что выражение r^2 характеризует долю систематических отклонений y в связи с изменением x, а разность ($1-r^2$) – долю случайных отклонений.

Результаты и их обсуждение

Результаты вычислений представлены в таблице 2. Приведены также промежуточные данные, являющиеся результатом поэтапного вычисления коэффициента корреляции. Для полной корреляционной решетки, полученной в эксперименте, коэффициент корреляции составил величину, равную $r=0,58$. Это дает все основания признать наличие корреляции между исследуемыми признаками. При этом важно отметить следующее. Корреляционная связь считается существенной, если прогнозируемая часть (r^2) составляет не менее 50% от общего изменения, т.е. коэффициент корреляции не менее 0,7. В нашем эксперименте приблизительно в 34% ($0,58^2 = 0,34$) случаев изменения y связаны с изменениями x, а остальные отклонения носят случайный характер (с точки зрения рассматриваемой связи).

Корреляционная решетка обладает большой наглядностью при анализе связи параметров. Однако при определении коэффициента корреляции ее построение необязательно. Коэффициент корреляции может быть вычислен и по несгруппированным данным. Пусть у каждого i-больного животного имеются измеренные значения параметров x_i и y_i. Тогда средние значения и средние квадратичные отклонения вычисляются по известным формулам:

$$\bar{x} = \frac{1}{N}\sum_{i=1}^{N} x_i, \quad (5)$$

$$\bar{y} = \frac{1}{N}\sum_{j=1}^{N} y_i, \quad (6)$$

$$\sigma_x = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})^2}, \quad (7)$$

$$\sigma_y = \sqrt{\frac{1}{N}\sum_{j=1}^{N}(y_j-\bar{y})^2}. \quad (8)$$

В этих формулах i-номер больного животного, N - общее число исследуемых животных. Момент ковариации в этом случае может быть найден из выражения

$$cov\{x,y\} = \frac{1}{N}\sum_{i=1}^{N}(x_i-x)(y_j-\bar{y}) = \frac{1}{N}\sum_{i=1}^{N} x_i y_i - \bar{x}\bar{y}. \quad (9)$$

После определения указанных величин коэффициент корреляции находится по формуле (4).

В заключение следует заметить, что высокие значения коэффициента корреляции ($r>0,8$) свидетельствуют о наличии тесной связи исследуемых параметров. При коэффициентах корреляции меньше 0,5 зависимость между параметрами следует признать незначительной. Необходимо, однако, сделать следующее пояснение. В некоторых случаях и только при нелинейной зависимости параметров значение r может получиться небольшим и при наличии явной корреляции. В подобных случаях нужно определять коэффициент корреляции для каждой области изменения параметров отдельно (например, разбив данные на две группы, содержащие показатели x меньше и больше среднего значения). Если и в этом случае коэффициенты корреляции получаются малыми, то следует сделать вывод об отсутствии корреляции признаков.

Таблица 2. Результаты поэтапных вычислений коэффициента корреляции для температуры и пульса

$\bar{x}$	$\bar{y}$	$\frac{1}{N}\sum_{i=1}^{N}(x_i-\bar{x})$	$\frac{1}{N}\sum_{j=1}^{N}(y_i-\bar{y})$	σ_x	σ_y	cov(x, y)	r
39,39	120,22	0,18	130,08	0,43	11,40	2,84	0,58