

### 4.3 Тема. Вычисление коэффициента корреляции при альтернативной изменчивости

**Цель.** Знакомство с методами вычисления коэффициента корреляции при альтернативной изменчивости.

Рассмотренные выше способы вычисления коэффициента корреляции употребляются в отношении признаков, сильно варьирующих и состояние которых можно измерить, взвесить и выразить в определенных цифровых показателях. Однако бывают случаи, касающиеся качественных признаков, когда требуется лишь констатировать наличие или отсутствие данного признака, или когда признак может проявиться лишь в двух состояниях. Такое положение называется альтернативным (двойко возможным), а признаки – альтернативными. Примером альтернативных признаков может служить наличие или отсутствие рогов у животных, той или иной окраски и т.д.

Корреляция между такими альтернативными признаками определяется степенью сочетаемости каждого из двух вариантов признака у двух взаимосвязанных групп особей.

Пример. Требуется установить связь между мастью матерей и мастью их потомства. Допустим от 300 красных матерей родилось 260 потомков красных и 40 палевых, а от 200 палевых матерей – 50 красных и 150 палевых. Эти данные можно разместить в простой корреляционной решетке (таблица 4.3.1).

Таблица 4.3.1 Распределение коров-матерей и их потомков в зависимости от масти

Масть потомков	Масть матерей		Итого
	красная	палевая	
Красная	260	50	310
Палевая	40	150	190
Итого	300	200	500

Обозначим четыре частоты, оказавшиеся в клетках нашей корреляционной решетки, через  $p_1, p_2, p_3, p_4$  (таблица 4.3.2).

Таблица 4.3.2 Вычисление коэффициента корреляции между альтернативными признаками

Масть потомков	Масть матерей		Итого
	1	2	
(1)	$P_1$	$P_2$	$(P_1 + P_2)$

(2)	P <sub>3</sub>	P <sub>4</sub>	(P <sub>3</sub> + P <sub>4</sub> )
Итого	(P <sub>1</sub> + P <sub>3</sub> )	(P <sub>2</sub> + P <sub>4</sub> )	N

Коэффициент корреляции (r) вычисляется по формуле:

$$r = \frac{P_1P_4 - P_2P_3}{\sqrt{(p_1 + p_2) \cdot (p_3 + p_4) \cdot (p_1 + p_3) \cdot (p_2 + p_4)}} \quad (43)$$

В нашем примере коэффициент корреляции между мастью матерей и мастью потомков равняется +0,62.

Это указывает, что большинство потомков имеет ту же масть, что и их матери.

Бисериальный коэффициент корреляции ( $r_{bs}$ ) применяется при измерении тесноты связи между качественными признаками, группируемыми в альтернативные группы (+ и -), и непрерывно варьирующими количественными признаками. Он вычисляется по формуле:

$$r_{(bs)} = \frac{\bar{x}_1 - \bar{x}_2}{\sigma_x} \sqrt{\frac{n_1 n_2}{N(N-1)}}, \quad (44)$$

где  $\bar{x}_1$  и  $\bar{x}_2$  - средние арифметические альтернативных групп,  $n_1$  и  $n_2$  - объемы этих групп;  $N=(n_1+n_2)$  - общее число наблюдений, или объем выборки;  $\sigma_x$  - среднее квадратическое отклонение для всех выборки.

Пример. Выяснилась значимость между полом подростков 16-17-летнего возраста и их тактильной чувствительностью. Единицей измерения признака служило расстояние между ножками эстеziометра (мм), при котором ощущение двух прикосновений к концу среднего пальца левой руки воспринималось как одно прикосновение, т.е. сливалось. Результаты опыта и их обработка приведены в таблице 4.3.3.

Сначала определяем средние арифметические групп:  $\overline{X_M} = 2,17$  мм,  $\overline{X_{Ж}} = 1,76$  мм. Затем находим величину среднего квадратического отклонения  $\sigma_x = 0,36$ . Подставляем найденные значения в формулу:

$$r_{bs} = \frac{2.17 - 1.76}{0.36} \sqrt{\frac{13 \cdot 15}{28 \cdot 27}} = \frac{0.41}{0.36} \sqrt{0.258} = \frac{0.146}{0.36} = 0.406.$$

Значимость выборочного  $r_{bs}$  оценивается с помощью t-критерия Стьюдента. С вероятностью  $P > 0,95$  можно заключить, что тактильная чувствительность конца среднего пальца левой руки у девушек выше, чем у юношей того же возраста.

Таблица 4.3.3

Показания эстеziометра, мм (x)	Пол подростков (y)		P <sub>i</sub>	P <sub>x<sub>i</sub></sub>	P <sub>x<sub>i</sub></sub> <sup>2</sup>
	мужской	женский			
1.5	1	5	6	9.0	13.50
1.6	-	2	2	3.2	5.12
1.7	-	2	2	3.4	5.78

1.8	1	1	2	3.6	6.48
1.9	1	1	2	3.8	7.22
2.0	2	2	4	8.0	16.00
2.1	-	1	1	2.1	4.41
2.2	1	-	1	2.2	4.84
2.3	2	-	2	4.6	10.58
2.4	3	-	3	7.2	17.28
2.5	2	1	3	7.5	18.75
$\Sigma$	$n_1=13$	$N_2=15$	$N=28$	54.6	109.96

**Задание 1.** От скрещивания самцов плодовой мушки дрозофилы, имеющих черную окраску тела и зачаточные крылья (рецессивные признаки), с нормальными самками того же вида, гетерозиготными по генам этих признаков, в потомстве оказались мухи:

Серые с нормальными крыльями...75

Серые с зачаточными крыльями...16

Черные с нормальными крыльями..14

Черные с зачаточными крыльями...68

Выясните, имеется ли связь между окраской тела и развитием крыльев у дрозофилы.

**Задание 2.** При изучении влияния отселекционированности на резистентность цыплят к пуллорозу оказалось, что после заражения живой культурой из 220 цыплят выжило 115, пало 105, а в отселекционированной группе выжило 560, пало 58.

Определить тетрахорический показатель связи между резистентностью цыплят к пуллорозу и степенью отселекционированности стада по этому показателю.

### **Контрольные вопросы.**

1. Как определить связь между качественными признаками?
2. В каких случаях возникает необходимость определения полихорического показателя связи, а в каких – бисериального?
3. Как вычисляется коэффициент корреляции для альтернативных признаков?