

Тема 4. ВЫЧИСЛЕНИЕ И АНАЛИЗ СТАТИСТИЧЕСКИХ ХАРАКТЕРИСТИК КАЧЕСТВЕННОЙ ИЗМЕНЧИВОСТИ

Объяснение. Основными статистическими показателями качественной изменчивости являются: доля признака, показатель изменчивости, коэффициент вариации и ошибка выборочной доли.

Доля признака, или относительная численность (частота) отдельной варианты в данной совокупности обозначается через P_1, P_2, P_3 и т.д. и может быть выражена в частях единицы или в процентах. В первом случае сумма всех долей в пределах данной совокупности или ряда распределения равна единице, а во втором – 100%.

Доля признака – это отношение численности каждого из членов ряда n_1, n_2, n_3 и т.д. к численности совокупности n , т.е. вероятность появления данного признака в изучаемой совокупности:

$$p_1 = \frac{n_1}{n}; p_2 = \frac{n_2}{n}; p_3 = \frac{n_3}{n} \text{ и т.д.}$$

При альтернативной (двоиковозможной) изменчивости доля одного признака обозначается через p , а второго через q . На основании очевидного равенства $p + q = 1$ (или 100%), так как сумма вероятностей двух противоположных событий всегда равна единице (100%), значение $q = 1 - p$.

Показатель изменчивости качественного признака S характеризует варьирование величин ряда относительно друг друга. Значение показателя изменчивости определяется по формуле:

$$S = \sqrt[k]{p_1 \times p_2 \times p_3 \times \dots \times p_k}$$

где p_1, p_2 и т.д. – доли признака (или процентные значения их) в общей совокупности;

k – число градаций признака.

Если изучаемая совокупность представлена объектами с двумя градациями признака (альтернативная изменчивость), то

$$S = \sqrt{pq}$$

где p и q – доли признака, выраженные в частях единицы или процентах.

Например, показатель изменчивости при $p = 0,10$ и $q = 0,90$ будет равен:

$$S = \sqrt{pq} = \sqrt{0,10 \times 0,90} = 0,30 \text{ (или 30\%)}$$

В зависимости от соотношения p и q значение S изменяется от 0 до 0,5. Максимальная изменчивость качественного признака S_{max} будет наблюдаться тогда, когда $p = q = 0,5$ и $S_{max} = \sqrt{0,5 \times 0,5} =$

0,5 (или 50%). Значения максимальной (наибольшей) изменчивости для распределений с разным числом градаций качественных признаков даны ниже:

Число градаций признака	S_{max}	Число градаций признака	S_{max}
2	0,500 (50,0%)	5	0,200 (20,0%)
3	0,333 (33,3%)	6	0,167 (16,7%)
4	0,250 (25,0%)	7	0,143 (14,3%)

Пользуясь величинами максимальных значений S_{max} , можно вычислить *коэффициент вариации качественных признаков* – фактический показатель изменчивости, выраженный в процентах к максимально возможной изменчивости:

$$V_p = \frac{S}{S_{max}} \cdot 100$$

Коэффициент вариации характеризует относительную степень изменчивости изучаемых признаков и широко используется для сравнительной оценки выравненности различных совокупностей. Максимальное значение $V_p = 100\%$ наблюдается при $S = S_{max}$.

Ошибка выборочной доли S_p – мера отклонения доли признака выборочной совокупности p от доли его по всей генеральной совокупности P вследствие неполной представительности (репрезентативности) выборки. Ошибку доли вычисляют по формуле:

$$S_p = \frac{S}{\sqrt{n}}$$

где S – показатель изменчивости качественного признака;
 n – объем выборки.

Для альтернативного варьирования, когда значение $S = \sqrt{pq}$, формула ошибки выборочной доли примет вид:

$$S_p = \sqrt{\frac{pq}{n}}$$

Вероятность встретить p (или q) в интервале $p \pm S_p$ составляет около 68%, в интервале $p \pm 2S_p$ – 95% и в интервале $p \pm 3S_p$ – около 99%. Следовательно, подобно количественной изменчивости, все значения p с вероятностью 99% укладываются в пределах тройной ошибки выборочной доли. Далее следует выполнить индивидуальное **задание** (приложение 6).