

6.2 Тема. Вычисление дисперсии однофакторного комплекса при большой выборке

Цель. Знакомство с методами вычисления дисперсии однофакторного комплекса при многочисленной выборке

Для вычисления общей дисперсии при многочисленной выборке пользуются следующей формулой:

$$C_y = \sum p_v a_v^2 - H, \text{ где } H = \frac{(\sum p_v a_v)^2}{n}.$$

Формула факториальной дисперсии: $C_x = \sum h_x - H.$

Формула остаточной дисперсии: $C_z = \sum p_v a_v^2 - \sum h_x.$

Пример. Влияние живой массы матерей на живую массу телят при рождении.

Порядок вычисления и необходимые данные приведены в таблице 6.2.1.

Таблица 6.2.1 Обработка однофакторного комплекса при многочисленной выборке

V (живая масса телят, кг)	Фактор А, живая масса матерей, кг				p _v	a _v	p _v a _v	p _v a _v ²
	400-449	450-499	500-549	550-599				
30-32,9	3	1	-	-	4	-3	-12	36
33-35,9	14	7	2	1	24	-2	-48	96
36-38,9	15	15	8	2	40	-1	40	40
39-41,9	8	10	14	14	46	0	0	0
42-44,9	2	5	10	16	33	1	33	33
45-47,9	-	2	2	7	11	2	22	44
48-50,9	-	-	-	2	2	3	6	18
P _A	42	40	36	42	160		-39	267
P _A a _v	-50	-23	2	32				
∑(p _A a _v) ²	2500	529	4	1024				
h _x =	59,5	13,2	0,1	24,4	97,2			
$\frac{(\sum p_A a_v)^2}{P_A}$								

Вычисленные выражения $\sum p_A a_v$ приводятся для каждого столбца по классам А. Для получения указанных выражений требуется умножить частоты по клеткам каждого столбца на отклонение a_v; сумму этих произведений записывают по каждому классу в строчку под строчкой p_A.

1 столбец:

$$\sum p_A a_v = 3(-3) + 14(-2) + 15(-1) + 8 \cdot 0 + 2 \cdot 1 = -50;$$

2 столбец:

$$\sum p_A a_v = 1(-3) + 7(-2) + 15(-1) + 10 \cdot 0 + 5 \cdot 1 + 2 \cdot 2 = -23;$$

3 столбец:

$$\sum p_A a_v = 2(-2) + 8(-1) + 14 \cdot 0 + 10 \cdot 1 + 2 \cdot 2 = 2;$$

4 столбец:

$$\sum p_A a_v = 1(-2) + 2(-1) + 14 \cdot 0 + 16 \cdot 1 + 7 \cdot 2 + 2 \cdot 3 = 32.$$

Промежуточная величина H равняется:

$$H = \frac{(\sum p_A a_v)^2}{n} = \frac{-39^2}{160} = 9,5.$$

Дисперсии C_y , C_x , C_z вычисляются по вышеприведенным формулам, подставляя в них данные из таблицы:

$$C_y = \sum p_v a_v^2 - H = 267 - 9,5 = 257,5;$$

$$C_x = \sum h_x - H = 97,2 - 9,5 = 87,7;$$

$$C_z = \sum p_v a_v^2 - \sum h_x = 267 - 97,2 = 169,8.$$

Доля учтенных факторов равняется:

$$\eta^2_x = \frac{87,7}{257,5} = 0,341, \text{ или } 34,1 \%$$

Доля неучтенных факторов равняется:

$$\eta^2_z = \frac{169,8}{257,5} = 0,659, \text{ или } 65,9 \%$$

Число степеней свободы для факториальной дисперсии равно:

$$v_x = l_x - 1 = 4 - 1 = 3.$$

Для остаточной дисперсии число степеней свободы равно:

$$v_z = n - l_x = 160 - 4 = 156.$$

Факториальная девиата равна: $\sigma^2_x = \frac{C_x}{v_x}$; в нашем примере $\sigma^2_x = \frac{87,7}{3} = 29,2$.

Остаточная девиата равна: $\sigma^2_z = \frac{C_z}{v_z}$; в нашем примере $\sigma^2_z = \frac{169,8}{156} = 1,2$.

Коэффициент достоверности Фишера равняется:

$$F = \frac{\sigma^2_x}{\sigma^2_z}; \text{ в нашем примере } F = 26,5.$$

Вычисленное значение F сравнивают с табличным значением F . Табличное значение F для данного примера при трех уровнях вероятности равно:

$$F_{0,95} = 2,7; F_{0,99} = 3,9; F_{0,999} = 5,7.$$

Влияние живой массы матерей на живую массу телят при рождении вполне достоверно при всех градациях вероятности, так как вычисленное F равно 26,5.

Задание. Влияние длительности сервис-периода на удой коров.

Варьирующий признак (длительность сервис-периода)	Фактор А – удой (кг) с градациями				Σ
	3000-3999	4000-4999	5000-5999	6000-6999	
18-18,9	3	10	6	-	
19-19,9	5	14	11	2	
20-20,9	4	22	20	5	
21-21,9	2	20	15	4	
22-23,9	1	8	9	5	

Обработайте эти данные методом дисперсионного анализа

Контрольные вопросы.

1. Когда необходим дисперсионный анализ?
2. Приведите формулы общей, факториальной и остаточной дисперсий.
3. Дайте определение равномерного однофакторного комплекса.