

Практическое занятие № 15

Вычисление точечной оценки параметров распределения по выборке и точечной несмещенной оценки для дисперсии.

Цель: научиться вычислять числовые характеристики вариационного ряда.

Норма времени: 2 часа.

Оборудование: тетрадь, ручка.

Порядок выполнения работы.

Краткая теория

Определение. Статистической оценкой Q^* неизвестного параметра Q теоретического распределения называют функцию $f(x_1, x_2, \dots, x_n)$ от наблюдаемых случайных значений x_1, x_2, \dots, x_n .

Определение. Точечной оценкой называют статистическую оценку, которая определяется одним числом $Q^* = f(x_1, x_2, \dots, x_n)$, где x_1, x_2, \dots, x_n — результаты n наблюдений над количественным признаком X (выборка).

Определение. Несмещенной называют точечную оценку Q^* , математическое ожидание которой равно оцениваемому параметру Q при любом объеме выборки, то есть $M(Q^*) = Q$. Смещенной называют точечную оценку, математическое ожидание которой не равно оцениваемому параметру.

Выборочная средняя.

Пусть для изучения генеральной совокупности относительно количественного признака X извлечена выборка объема n .

Определение. Выборочной средней \bar{x}_B называют среднее арифметическое значение признака выборочной совокупности.

Если все значения x_1, x_2, \dots, x_n признака выборки объема n различны, то

$$\bar{x}_B = (x_1 + x_2 + \dots + x_n) / n$$

Если же все значения признака x_1, x_2, \dots, x_k имеют соответственно частоты n_1, n_2, \dots, n_k , причем $n_1 + n_2 + \dots + n_k = n$, то

$$\bar{x}_B = \left(\sum_{i=1}^k n_i \cdot x_i \right) / n,$$

где $n = \sum_{i=1}^k n_i$ — объем выборки.

Выборочная средняя является несмещенной оценкой генеральной средней (неизвестного математического ожидания).

Замечание. Если первоначальные варианты x_i — большие числа, то для упрощения решения целесообразно вычесть из каждой варианты одно и то же число C , то есть перейти к условным вариантам $u_i = x_i - c$. Тогда

$$\bar{x}_B = c + \left(\sum_{i=1}^k n_i \cdot u_i \right) / n.$$

Выборочная дисперсия.

Смещенной оценкой генеральной дисперсии служит выборочная дисперсия. Эту величину вводят для того, чтобы охарактеризовать рассеяние наблюдаемых значений количественного признака выборки вокруг среднего значения \bar{x}_B .

Определение. Выборочной дисперсией D_B называют среднее арифметическое квадратов отклонения наблюдаемых значений признака от их среднего значения \bar{x}_B . Если значения признака x_1, x_2, \dots, x_k имеют соответственно частоты n_1, n_2, \dots, n_k , причем $n_1 + n_2 + \dots + n_k = n$, то

$$D_B = \left(\sum_{i=1}^k n_i (x_i - \bar{x}_B)^2 \right) / n.$$

Эта оценка является смещенной, так как $M(D_B) = \frac{n-1}{n} \cdot D_\Gamma$,

где D_Γ – генеральная дисперсия – среднее арифметическое квадратов отклонения значения признака генеральной совокупности от их среднего значения \bar{x}_n .

Теорема. Выборочная дисперсия равна среднему квадратов значений признака минус квадрат выборочной средней.

$$D_B = \overline{x^2} - [\bar{x}]^2 = \frac{\sum n_i x_i^2}{n} - \left[\frac{\sum n_i x_i}{n} \right]^2.$$

Для вычисления выборочной дисперсии эта формула наиболее удобна.

Замечание. Если перейти к условным вариантам $u_i = x_i - c$, то дисперсия при этом не

изменится. Тогда $D_B(x) = D_B(u) = \overline{u^2} - [\bar{u}]^2 = \frac{\sum n_i \cdot u_i^2}{n} - \left[\frac{\sum n_i \cdot u_i}{n} \right]^2.$

Оценка генеральной дисперсии по исправленной выборочной.

Пусть из генеральной совокупности в результате n независимых наблюдений над количественным признаком X извлечена повторная выборка объема n :

Значения признака	x_i	x_1	x_2	...	x_k
Частоты	n_i	n_1	n_2	...	n_k

При этом $n_1 + n_2 + \dots + n_k = n$. Требуется по данным выборки найти неизвестную генеральную дисперсию D_Γ . Если в качестве оценки D_Γ принять выборочную дисперсию, то эта оценка будет приводить к систематическим ошибкам, давая заниженное значение D_Γ . Объясняется это тем, что математическое ожидание выборочной дисперсии не равно оцениваемой D_Γ , а равно $M[D_B] = \frac{n-1}{n} D_\Gamma$.

Легко «исправить» выборочную дисперсию так, чтобы ее математическое ожидание было равно генеральной дисперсии. Достаточно для этого умножить D_B на дробь $n/(n-1)$. Сделав это, мы получим *исправленную дисперсию*, которую обычно обозначают S^2 .

Несмещенной оценкой генеральной дисперсии служит исправленная выборочная дисперсия:

$$S^2 = \frac{n}{n-1} D_B = \frac{\sum n_i (x_i - \bar{x}_B)^2}{n-1}.$$

Более удобна форма:

$$S_x^2 = \frac{\sum n_i x_i^2 - [\sum n_i x_i]^2 / n}{n-1}.$$

В условных вариантах она имеет вид:

$$S_u^2 = \frac{\sum n_i u_i^2 - [\sum n_i u_i]^2 / n}{n - 1},$$

причем если $u_i = x_i - c$, то $S_x^2 = S_u^2$; если $u_i = c \cdot x_i$, то $S_x^2 = S_u^2 / c^2$.

Задача 1.

Из генеральной совокупности извлечена выборка объемом $n = 60$

x_i	1	3	6	26
n_i	8	40	10	2

Найти несмещенную оценку генеральной средней.

Решение. Несмещенной оценкой генеральной средней является выборочная

$$\bar{x}_B = \left(\sum_{i=1}^k n_i \cdot x_i \right) / n,$$

где x_i — варианты выборки, n_i — частота варианты x_i ; $n = \sum_{i=1}^k n_i$ — объем выборки.

$$\bar{x}_B = (8 \cdot 1 + 40 \cdot 3 + 10 \cdot 6 + 2 \cdot 26) / 60 = (8 + 120 + 60 + 52) / 60 = 240 / 60 = 4.$$

Ответ: $\bar{x}_B = 4$.

Задача 2.

Выборочная совокупность задана таблицей распределения

x_i	1	5	3	4
n_i	20	15	10	5

Найти выборочную дисперсию.

Решение. Найдем выборочную среднюю

$$\bar{x}_B = \left(\sum_{i=1}^4 n_i \cdot x_i \right) / n = \frac{20 \cdot 1 + 15 \cdot 2 + 10 \cdot 3 + 5 \cdot 4}{20 + 15 + 10 + 5} = \frac{100}{50} = 2.$$

Найдем выборочную дисперсию:

$$D_B = \left(\sum_{i=1}^4 n_i (x_i - \bar{x}_B)^2 \right) / n,$$

$$D_B = \frac{20 \cdot (1 - 2)^2 + 15 \cdot (2 - 2)^2 + 10 \cdot (3 - 2)^2 + 5 \cdot (4 - 2)^2}{50} =$$

$$= \frac{20 \cdot (-1)^2 + 15 \cdot 0 + 10 \cdot 1^2 + 5 \cdot 2^2}{50} = \frac{20 + 10 + 20}{50} = \frac{50}{50} = 1.$$

Ответ: $D_B = 1$.

Задания для выполнения:

1 Вариант

1. По выборке объема $n=81$ найдена смещенная оценка $D_B=5$ генеральной дисперсии. Найти несмещенную оценку дисперсии генеральной совокупности.
2. Найти выборочную дисперсию по данному распределению выборки объема $n = 27$:

x_i	354	365	372
n_i	4	9	14

3. Найти выборочную дисперсию по данному распределению выборки объема $n = 120$:

x_i	3832	3848	3850	3900
n_i	13	24	35	48

4. Найти выборочную дисперсию по данному распределению выборки объема $n = 34$:

x_i	344	349	355
n_i	6	8	20

2 Вариант

1. В итоге пяти измерений (без систематических ошибок) длины бруска одним прибором получены следующие результаты: 804, 806, 807, 809, 810. Найти: а) выборочную среднюю длину бруска; б) выборочную и исправленную дисперсии ошибок прибора.

2. Найти выборочную дисперсию по данному распределению выборки объема $n = 75$:

x_i	34,7	35,4	35,9	36,3
n_i	13	18	24	20

3. Найти выборочную дисперсию по данному распределению выборки объема $n = 20$:

x_i	0,004	0,005	0,008
n_i	4	7	9

4. Найти выборочную дисперсию по данному распределению выборки объема $n = 44$:

x_i	0,3	0,7	0,9
n_i	7	15	22

Контрольные вопросы:

1. В чем сущность задачи по определению параметров генеральной совокупности?
2. Какую величину принимают за среднюю генеральной совокупности? Как она вычисляется?
3. Какую величину принимают за дисперсию генеральной совокупности? Как она вычисляется?