Тема лекции 5.

Статистические оценки параметров распределения случайных величин по выборкам. Степенные средние.

Степенные средние и способы их вычисления.
Средняя взвешенная
Средняя геометрическая
Средняя квадратичная
Средняя кубическая
Средняя гармоническая
Средняя арифметическая для альтернативных признаков.

Степенные средние и способы их вычисления. *Средняя арифметическая* ($\overline{X}$ – для выборочной совокупности; $\tilde{X}$ – для генеральной совокупности). Этот показатель является центром распределения, вокруг которого группируются все варианты статистической совокупности. Средняя арифметическая может быть простой ($\overline{X}$) и взвешенной ($\overline{X}_{\text{взв}}$).

Методы вычисления средней арифметической ($\overline{X}$):

а) прямой, когда среднюю арифметическую определяют как сумму вариант всех членов совокупности, деленную на их число:

$$\overline{X} = \frac{x_1 + x_2 + \ldots + x_n}{n} = \frac{\sum x_n}{n} \qquad (5)$$

прямой метод вычисления $\overline{X}$ приемлем для малых выборок.

б) непрямой, приемлемый для больших выборок:

$$\overline{X} = \mathrm{A} + \mathrm{k}\frac{\sum fa}{n} \qquad (6)$$

где А – условная средняя, взятая из модального класса.

Средняя взвешенная ($\bar{X}_{вз}$) – это результат усреднения средних арифметических нескольких совокупностей. Для вычисления этого показателю пользуются формулой:

$$\bar{X}_{вз} = \frac{\bar{X}_1 \cdot n_1 + \bar{X}_2 \cdot n_2 + \ldots + \bar{X}_n \cdot n_n}{n_1 + n_2 + \ldots + n_n} = \frac{\sum \bar{X} \cdot n}{\sum n} \qquad (7)$$

Условие. Имеются средние арифметические показатели длины тела (в см) отдельных лабораторных крыс и их количество в разных лабораториях:

№ стад	1	2	3	4	5
$\overline{X}$	20	22	25	18	24
n	210	150	240	100	300

Требуется вычислить среднюю длину тела лабораторных крыс по этим лабораториям.

Решение. В данном случае нельзя вычислять среднюю арифметическую путем сложения всех средних по длине тела и деления суммы на количество лабораторий, так как средняя длина тела в каждой отдельной лаборатории относится к разному количеству крыс. Поэтому для вычисления пользуются формулой 6:

$$X_{взв} = \frac{20 \cdot 210 + 22 \cdot 150 + 25 \cdot 240 + 18 \cdot 100 + 24 \cdot 300}{210 + 150 + 240 + 100 + 300} = \frac{4200 + 3300 + 6000 + 1800 + 7200}{1000} = \frac{22500}{1000} = 22{,}5$$

Ответ. Средняя длина тела крыс по пяти лабораториям составляет:

$X_{взв} = 22{,}5$ см.

Средняя геометрическая ($\overline{X}_g$) – это среднее значение признака, характеризующий темп роста, темп увеличения популяции. Применение данного показателя удобно в тех случаях, когда признак выражен в долях единицы или в процентах и изменяется во времени или по периодам. При вычислении $\overline{X}_g$ необходимо исключать варианты, выражающиеся нулем или отрицательным числом. С помощью $\overline{X}_g$ можно определить относительный прирост стада, привесов тела за определенный период времени. Формула для вычисления средней геометрической:

$$\overline{X}_g = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n} \qquad (8)$$

где, x – варьирующий признак,

n – число наблюдений в выборке.

Условие. Допустим, необходимо определить темп увеличения численности популяции бактерии кишечной палочки *E.coli* за 7 мин: 5-8-25.

Решение. Для выполнения этого задания необходимо воспользоваться формулой 8.

$$\overline{X}_g = \sqrt[n]{x_1 \cdot x_2 \cdot \ldots \cdot x_n} = \sqrt[3]{5 \cdot 8 \cdot 25} = \sqrt[3]{1000} = 10$$

Ответ. За 7 минут, в среднем, популяция бактерии E.coli увеличивается на 10.

Средняя квадратичная ($\overline{X}_s$) – характеристика мер площади:

$$\overline{X}_s = \sqrt{\frac{\sum x^2}{n}} \qquad (9)$$

где $\sum x^2$ - сумма квадратов варьирующего признака,

n – число наблюдений в выборке.

Условие. Необходимо определить средний диаметр ядра в клетках следующего вариационного ряда (x – диаметр клеток, микрон):

№	x	x^2	№	x	x^2
1	15	225	7	20	400
2	12	144	8	19	361
3	20	400	9	17	289
4	22	484	10	14	196
5	18	324	11	14	196
6	15	225	12	12	144
					$\sum x^2 = 3388$

Решение. Используя формулу 9, находим

$$\overline{X}_s = \sqrt{\frac{\sum x^2}{n}} = \sqrt{\frac{3388}{12}} = \sqrt{282,33} = 16,80$$

Ответ. Средний диаметр ядер равен 16,80 μ.

Средняя кубическая ($\overline{X}_Q$) – характеристика объемных признаков:

$$\overline{X}_Q = \sqrt[3]{\frac{\sum fx^3}{n}} \qquad (10)$$

Условие. Допустим, необходимо определить средний объем яиц по их диаметрам. Объем 18 куриных яиц (учитывали полусумму большого и малого диаметра) были следующими (диаметр яиц – x, см; число случаев – f):

x, см	4,7	4,8	5,0	5,4	5,6	6,0	
f	2	4	6	3	2	1	Σf=n=18

Решение. Средний объем яиц при использовании формулы 9 составит:

$$\overline{X}_Q = \sqrt[3]{\frac{\sum fx^3}{n}} =$$

$$\sqrt[3]{\frac{2\cdot4{,}7^3+4\cdot4{,}8^3+6\cdot5{,}0^3+3\cdot5{,}4^3{,}2\cdot5{,}6^3+1\cdot6{,}0^3}{18}} = \sqrt[3]{\frac{2419{,}638}{18}} = \sqrt[3]{134{,}42} = 5{,}12\,см$$

Ответ. Средний объем яиц равен 5,12 см.

Средняя гармоническая ($\overline{X}_{\text{н}}$) – показатель используется для определения средних изменяющихся скоростей. Для таких процессов характерно, что при увеличении одного показателя другой изменяется в обратном направлении, т.е. уменьшается. Например, чем быстрее бежит лошадь, тем меньше она тратит время на прохождение пути.

$$\overline{X}_{\text{н}} = \frac{n}{\sum \frac{1}{x}} \qquad (11)$$

В заключении обзора степенных средних необходимо отметить, что между ними существуют определенные соотношения, выражаемые следующим рядом мажорантности (неравенства):

$$\overline{X}_{\text{g}} > \overline{X}_{\text{s}} > \overline{X} > \overline{X}_{\text{Q}} > \overline{X}_{\text{H}}. \qquad (12)$$

Из этого ряда следует, что $\overline{X}_{\text{H}}$ всегда меньше, чем $\overline{X}$, Mo, Me и других средних величин.

Условие. Допустим, пять школьников возрастом 10 лет в течение 1 часа (60 мин) прочитали следующее количество страниц:

первый –10с; второй – 20с; третий – 25 с; четвертый – 30 с; пятый – 20 с. Всего - 105 с.

Решение. При использовании формулы 10 мы получим среднее количество страниц, прочитанное ими за 60 мин:

$$\overline{X}_{н} = \frac{n}{\sum \frac{1}{x}} = \frac{5}{\frac{1}{10}+\frac{1}{20}+\frac{1}{25}+\frac{1}{30}+\frac{1}{20}} = \frac{5}{0{,}273} = 18{,}31 стр$$

Ответ. Среднее количество страниц, прочитанное за 60 мин школьниками равно 18,31с.

Средняя арифметическая для альтернативных признаков. При вычислении средней арифметической для альтернативных признаков применяют показатель доли, в которую входят члены совокупности, имеющие данный альтернативный признак. Это можно выразить следующей формулой:

$$\overline{X}_{альт} = \frac{p}{n} \qquad (13)$$

где p- число членов совокупности с наличием альтернативного признака;

n – общее число членов выборки.

Показатель разнообразия для альтернативных признаков определяют при помощи среднего квадратического отклонения в абсолютных и относительных выражениях по формулам:

$$\sigma = \sqrt{p \cdot q} \qquad (14)$$

или $$\sigma = \sqrt{p \cdot (1 - q)} \qquad (15)$$

или $$\sigma = \sqrt{n \cdot p \cdot q} \qquad (16)$$

где: р- доля особей, имеющих данный признак в совокупности;

q- доля особей, лишенных данного признака.

Если требуется получить величину стандартного отклонения, выраженную в абсолютных показателях, необходимо воспользоваться формулой 15.

Необходимо помнить, что величина стандартного отклонения для качественных признаков не больше 0,5 или 50%.

Литература:

Основная – 1 [6-88]; 2 [т.1-11-78]; 3 [14-17].

Дополнительная – 3 [9-24]; 5 [8-16].

Контрольные вопросы:

1. Показатели средних величин: свойства и особенности их применения.
2. Структурные средние и способы их вычисления.
3. Степенные средние и способы их вычисления.
4. Средняя гармоническая.
5. Средняя арифметическая для альтернативных признаков.