

Практическое занятие № 5
СТАТИСТИЧЕСКАЯ ОБРАБОТКА РЕЗУЛЬТАТОВ
НАБЛЮДЕНИЙ ПРИ КОЛИЧЕСТВЕННОЙ
ИЗМЕНЧИВОСТИ

Количественная изменчивость – это такая изменчивость, при которой значение варьирующего признака имеет числовое выражение. Она бывает прерывистой (дискретной) и непрерывной. При *прерывистой* изменчивости значения признака выражаются только целыми числами; при *непрерывной* количественной изменчивости значения признака могут иметь любую величину в зависимости от точности, принимаемой для характеристики данного признака.

При проведении наблюдения исследователь получает ряд значений изучаемого варьирующего признака X , который подвергается статистической обработке. Основными статистическими характеристиками количественной изменчивости являются: *средняя арифметическая, дисперсия, стандартное отклонение, коэффициент вариации, ошибка выборочной средней и относительная ошибка выборочной средней.*

Средняя арифметическая \bar{x} – это обобщенная, абстрактная характеристика совокупности. Она не содержит полной информации о варьирующих объектах. При одинаковых средних характеризующие признаки могут отличаться по величине вариации.

Различают простую и взвешенную среднюю арифметическую. *Простая средняя арифметическая* рассчитывается для выборок малого объема по формуле:

$$\bar{x} = \frac{\sum X}{n} \quad (17)$$

где $\sum X$ – сумма всех значений признака,

n – количество значений признака, или объем выборки.

Взвешенная средняя арифметическая рассчитывается для сгруппированных данных по формуле:

$$\bar{x} = \frac{f_1 \cdot x_1 + f_2 \cdot x_2 + \dots + f_n \cdot x_n}{f_1 + f_2 + \dots + f_n} = \frac{\sum fX}{n} \quad (18)$$

где n – значение признака,

f – частота встречаемости каждого признака,

n – объем выборки.

Основное свойство средней арифметической заключается в равенстве суммы всех положительных и всех отрицательных отклонений от нее, т.е. сумма всех отклонений вариантов равна 0.

Дисперсия s^2 , или **варианса**, или **средний квадрат** рассчитывается как отношение суммы квадратов отклонений среднего арифметического от каждого значения признака к числу степеней свободы.

Для выборки малого объема она рассчитывается по формуле:

$$s^2 = \frac{\sum (X - x)^2}{n - 1} = \quad (19)$$

где $n-1$ – число степеней свободы, или количество варьирующих величин.

Для сгруппированных отклонений:

$$s^2 = \frac{\sum f \cdot (X - x)^2}{n - 1} = \quad (20)$$

Дисперсия показывает квадрат среднего отклонения значений признака X от средней арифметической.

Стандартное отклонение s , или среднее квадратическое отклонение получают путем извлечения квадратного корня из дисперсии, т.е. $s = \sqrt{s^2}$.

Стандартное отклонение показывает величину среднего отклонения значений признака X от средней арифметической.

Коэффициент вариации V – является относительным показателем изменчивости и представляет собой отношение стандартного отклонения к средней арифметической, выраженное в процентах:

$$V = \frac{s}{x} \cdot 100 = \quad (21)$$

Коэффициент вариации показывает *степень* изменчивости признака: изменчивость считается незначительной, если коэффициент вариации не превышает 10%, средним – если он колеблется от 10 до 20 и значительной – если он более 20%.

Коэффициента вариации используется только в том случае, если признак имеет только **положительные** значения.

Ошибка выборочной средней или ошибка выборки s_x является мерой отклонения выборочной средней арифметической от средней генеральной совокупности. Ошибки выборки возникают вследствие неполной репрезентативности (представительности), т.к. не все возможные значения признака, имеющиеся в генеральной совокупности, попадают в выборку. Она свойственна только выборочному методу. Величина ошибки выборочной средней зависит от степени варьирования и от объема выборки.

Ошибка выборочной средней прямо пропорциональна стандартному отклонению и обратно пропорциональна объему выборки, т.е.

$$s_x = \frac{s}{\sqrt{n}} = \quad (22)$$

Ошибка выборочной средней выражается в тех же единицах измерения, что и варьирующий признак и приписывается к соответствующим средним со знаком \pm , т.е. $\bar{x} \pm s_x$.

Пример. Масса зерен с одного растения ячменя сорта Донской, г.

Целью данного наблюдения является определение выравненности признака «масса зёрен с одного растения» у ячменя сорта Донской. Для достижения этой цели следует провести статистический анализ полученных данных:

10,5 12,4 11,0 11,1 11,6 12,4 11,4 13,2 11,4 10,8
 11,9 9,6 11,8 12,2 10,3 11,7 12,0 11,3 12,5 11,0
 12,3 12,0 10,4 11,5 10,0 10,8 10,6 11,8 11,6 11,1
 11,2 10,9 12,8 11,3 12,6 11,2 11,9 12,7 10,7 11,3

Объем выборки $n=40$, поэтому для проведения статистического анализа следует построить вариационный ряд. Определяем количество интервалов k и шаг интервала i :

$$k = \sqrt{n} = 6,3 \approx 6 ; \quad i = \frac{X_{\max} - X_{\min}}{k} = \frac{13,2 - 9,6}{6} = 0,6$$

Строим вариационный ряд и заполняем вспомогательную расчетную таблицу:

Интервал	Частота, f	X	$f \cdot X$	X^2	$f \cdot X^2$
9,6 ÷ 10,2	3	9,9	29,7	98,0	294,0
10,3 ÷ 10,8	6	10,6	63,6	112,4	674,4
10,9 ÷ 11,4	12	11,2	134,4	125,4	1504,8
11,5 ÷ 12,0	10	11,8	118,0	139,2	1392,0
12,1 ÷ 12,6	6	12,4	74,4	153,8	922,8
12,7 ÷ 13,2	3	13,0	39,0	169,0	507,0
Суммы	40	-	459,1	-	5295,0

Проводим статистическую обработку данных наблюдения:

Средняя арифметическая взвешенная

$$\bar{x} = (\sum f \cdot X) : n = 459,1 : 40 = 11,5$$

$$\begin{aligned} \text{Сумма квадратов отклонений } \sum (X - \bar{x})^2 &= \sum f \cdot X^2 - (\sum f \cdot X)^2 : n = \\ &= 5295,0 - 459,1^2 : 40 = 5295,0 - 21072,8 : 40 = 5295,0 - 5269,3 = 25,7 \end{aligned}$$

$$\text{Дисперсия } s^2 = \frac{\sum (X - \bar{x})^2}{n - 1} = \frac{25,7}{40 - 1} = \frac{25,7}{39} = 0,66$$

$$\text{Стандартное отклонение } s = \sqrt{s^2} = \sqrt{0,66} = 0,81$$

$$\text{Коэффициент вариации} \quad V = \frac{s}{\bar{x}} \cdot 100 = \frac{0,81}{11,5} \cdot 100 = 7,0$$

Значение коэффициента вариации меньше 10%, значит изменчивость признака незначительная, следовательно, признак «масса зерен с одного растения» у ячменя сорта Донской выровнен.

$$\text{Ошибка выборочной средней} \quad S_x = \frac{s}{\sqrt{n}} = \frac{0,81}{\sqrt{40}} = \frac{0,81}{6,32} = 0,1$$

Доверительный интервал для 95% уровня вероятности равен:

$$\bar{x} \pm t \cdot s_x \quad 11,5 \pm 2,03 \times 0,1 \quad 11,5 \pm 0,2 \quad \text{или} \quad 11,3 \div 11,7$$

Значения t_{05} берем из статистической таблицы (приложение 2) для $v=40-1=39$ степеней свободы. Для $v=39$ значение $t_{05} = 2,03$.

Следовательно, генеральная средняя признака «масса зерен с 1 растения» у ячменя сорта Донской с вероятностью 95% находится в интервале от 11,3 до 11,7 г. Вероятность того, что значение генеральной средней не попало в границы интервала, составляет 5%.

Задание. Согласно N варианта провести статистический анализ результатов наблюдения.

Шифр задания к работе

№ варианта	Номера колонок	№ варианта	Номера колонок	№ варианта	Номера колонок
1	1 2 3 4	6	1 3 5 7	11	3 4 7 8
2	2 3 4 5	7	2 4 6 8	12	2 3 6 7
3	3 4 5 6	8	3 5 7 9	33	3 4 7 8
4	4 5 6 7	9	4 6 8 10	14	5 6 9 10
5	7 8 9 10	10	1 2 5 6	15	1 4 7 10

Пример 1. Масса плодов тыквы сорта Чародейка, кг

Значение признака по колонкам

1	2	3	4	5	6	7	8	9	10
4,0	4,1	5,3	4,5	6,4	5,1	4,2	4,8	4,3	3,6
5,1	5,2	4,5	5,5	5,3	3,7	4,8	4,1	4,7	5,4
5,7	3,3	5,2	6,0	5,6	5,0	3,2	5,2	5,5	4,8
5,4	5,8	4,3	4,9	5,0	4,4	3,9	5,8	6,4	4,5
4,4	5,9	6,3	5,4	4,3	4,0	5,7	4,4	4,6	6,0
4,1	5,3	5,1	4,8	5,5	4,7	4,9	5,0	3,4	5,7
3,7	5,2	5,8	5,4	4,4	5,6	4,3	4,8	5,3	5,3
4,6	4,7	4,3	4,2	3,0	4,9	5,3	4,2	4,5	4,1
6,2	5,4	3,6	5,3	5,1	6,2	4,9	3,8	5,9	6,5
5,5	4,2	5,2	3,7	3,9	5,5	5,1	5,4	3,7	4,6

Пример 2. Содержание сахара в корнеплодах сахарной свеклы гибрида Ванесса, %

Значение признака по колонкам

1	2	3	4	5	6	7	8	9	10
13,9	16,6	15,7	14,2	15,5	13,2	16,1	13,1	16,3	18,7
15,1	15,7	13,5	15,3	18,9	15,6	15,4	15,5	17,8	15,1
14,5	12,1	16,7	17,4	19,9	17,5	12,2	14,3	16,0	14,5
16,0	16,4	14,8	16,2	15,0	15,4	15,3	15,7	13,5	17,0
13,1	17,8	14,0	13,7	15,4	18,6	14,9	17,3	14,7	16,4
15,6	15,2	15,8	14,9	16,1	16,5	14,2	15,4	15,9	18,0
18,0	13,3	14,3	12,8	17,2	16,6	13,3	14,8	12,8	16,8
14,6	16,4	15,6	15,5	16,3	14,1	16,5	16,9	15,2	14,6
18,7	14,4	17,6	16,8	13,8	13,7	15,1	18,4	14,4	15,8
16,5	14,7	19,0	16,2	14,1	15,0	17,7	16,2	16,7	13,9

Выполнение работы

1. Записать название примера, номер варианта
2. Заполнить рабочую таблицу

Номера колонок	Значение признака, X

3. Подсчитать количество значений признака $n =$
4. Определить количество интервалов $k = \sqrt{n} =$
5. Рассчитать величину интервала $i = \frac{X_{\max} - X_{\min}}{k} =$
6. Заполнить расчетную таблицу

Интервал	Разноска значений признака	Частота, f	X	$f \cdot X$	X^2	$f \cdot X^2$
Суммы	-		-		-	

7. Провести статистическую обработку данных

Средняя арифметическая $\bar{x} = (\sum f \cdot X) : n =$

Сумма квадратов отклонений $\sum (X - \bar{x})^2 = \sum f \cdot X^2 - (\sum f \cdot X)^2 : n =$

Дисперсия $s^2 = \frac{\sum (X - \bar{x})^2}{n - 1} =$

Стандартное отклонение $s = \sqrt{s^2} =$

Коэффициент вариации $V = \frac{s}{\bar{x}} \cdot 100 =$

Ошибка выборочной средней $S_x = \frac{s}{\sqrt{n}} =$

Доверительный интервал $\bar{x} \pm t_{05} \cdot S_x$

при $v = n - 1 =$

$t_{05} =$

Контрольные вопросы к работе:

1. Что такое наблюдение и эксперимент?
2. Что называется изменчивостью? Виды изменчивости
3. Какая изменчивость называется количественной? Виды количественной изменчивости
4. Основные статистические характеристики количественной изменчивости