

**Домашнее задание по курсу лекций
«Статистика». Часть 1
2016 год**

Тема: «Ряды распределения. Средние величины. Показатели вариации»
Работа выполняется студентами на листах формата А4.

Первый лист – титульный. На нем указывается тема работы, учебная группа, номер варианта студента, а также Ф.И.О. студента и преподавателя, принимающего работу.

ВАЖНО! В работе во всех задачах необходимо приводить **расчетные формулы** и отражать **ход расчетов**.

В завершение каждого пункта задачи необходимо привести выводы по результатам проведенных расчетов.

Домашнее задание содержит двадцать восемь вариантов. Выбор варианта определяется порядковым номером студента в журнале учета успеваемости.

Исходные данные, необходимые для выполнения работ, даны в Таблице №2, в соответствии со схемой по Таблице №1:

Таблица №1. Номера предприятий для варианта по журналу

Номер варианта по журналу	Номера предприятий
1	1-30
2	3-32
3	5-34
4	2-31
5	13-42
6	15-44
7	18-47
8	19-48
9	12-41
10	6-35
11	10-39
12	16-45
13	20-49
14	7-36
15	29-52,10-15

Номер варианта по журналу (продолжение)	Номера предприятий (продолжение)
16	17-46
17	23-52
18	22-51
19	9-38
20	21-50
21	8-37
22	4-33
23	11-40
24	19-48
25	14-43
26	22-51
27	13-42
28	15-44

Статистический ряд распределения - это упорядоченное распределение единиц совокупности на группы по определенному признаку (т.е. частный случай группировок вообще). Построив ряд, но и представить его графически в виде гистограммы, кумуляты, огивы, а также рассчитать средние величины: арифметическую и структурные – моду, медианы, квартили.

Исследование вариации является составным элементом статистического анализа, позволяющим оценить колебания значений изучаемого признака, однородность совокупности по данному признаку, взаимосвязь его с другими признаками. Показатели вариации служат критерием типичности рассчитанных по совокупности средних величин.

При **расчете основных показателей вариации** применяются **взвешенные** соответствующие показатели.

Следует помнить, что расчет показателей не самоцель статистического исследования, поэтому необходимо полученные величины проанализировать.

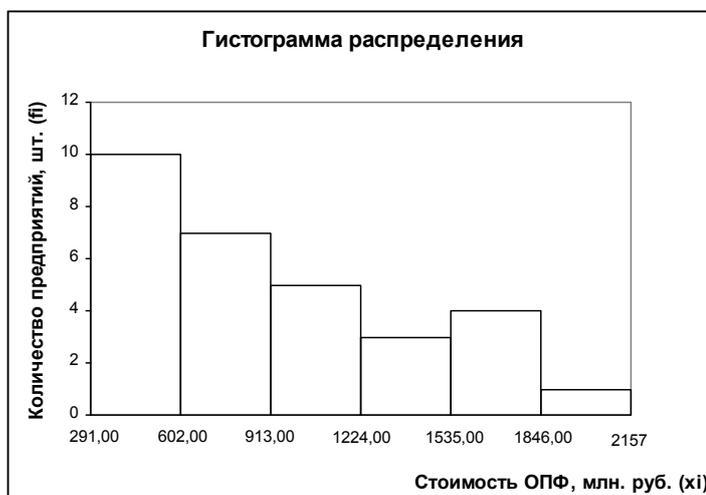
Задача 1. Построение ряда распределения

1.1. По исходным данным, представленным в **Таблице №1** (Вашего варианта), постройте **ряд распределения по Стоимости основных производственных фондов (ОПФ)** предприятий, образовав группы предприятий с равными интервалами. Результаты представьте в табличной форме:

Группы по стоимости ОПФ, млн.руб. (x_i)	Количество предприятий (частота, f_i)	Накопленная частота (f_i^H)	Частость (p_i^H)	Накопленная частость (p_i^H) <i>Для кумуляты (p_i^H) или</i> <i>Для огивы (p_i^H-огива)</i>

1.2. По данным ряда распределения (см. пункт 1.1) **постройте:**

- для студентов с **нечетным номером варианта** по журналу - гистограмму и кумуляту распределения,
- для студентов с **четным номером варианта** по журналу - гистограмму и огиву.



Пример построения Гистограммы распределения

Сформулируйте краткие выводы.

Задача 2. Характеристика ряда распределения

2.1. По полученному в **Задаче 1** (интервальному) ряду распределения определите:

- а) среднюю стоимость ОПФ,
 - б) моду,
 - в) медиану,
 - г) верхний, нижний квартили, межквартильное расстояние,
 - д) коэффициенты: асимметрии и эксцесса
- сформулируйте выводы о средних значениях и форме распределения.

2.2. Рассчитайте по интервальному ряду **показатели разброса (вариации)**:

- а) среднее линейное отклонение,
- б) дисперсию, среднеквадратическое отклонение,
- в) относительное линейное отклонение,
- г) коэффициент вариации.

По результатам расчетов сформулировать **вывод** об однородности или неоднородности совокупности.

Вопросы для самопроверки.

1. Что представляют собой ряды распределения?
2. По каким признакам могут быть образованы ряды распределения?
3. Как подразделяются вариационные ряды, и на каких признаках основано такое деление?
4. Что такое плотность распределения, для чего она используется?
5. Что такое полигон и гистограмма, для чего они применяются и как строятся?
6. Как строятся кумюлята и огива, что они характеризуют?
7. Что такое частость ряда распределения?
8. Дать формулы средней арифметической простой и взвешенной.
9. Как изменится средняя, если размер варьирующего признака уменьшить на одно и то же число?
10. Как изменится средняя, если размер варьирующего признака увеличить в одно и то же число раз?
11. Как изменится средняя, если кратно увеличить все веса?
12. Чему равна сумма всех отклонений индивидуальных значений от средней арифметической?
13. Назовите основные структурные средние. Как они рассчитываются?
14. Почему размах вариации является не надежным показателем колеблемости?
15. Что такое дисперсия и как она вычисляется?
16. Что характеризует среднее линейное отклонение?
17. Какие выводы можно сделать на основе коэффициента вариации?
18. В чем смысл правила сложения дисперсий?
19. Как можно оценить правильность выбора группировочного признака?

Статистические материалы для выполнения домашнего задания

Таблица №2. Техничко-экономические показатели работы предприятий отрасли одного из регионов (цифры условные).

Номер предприятия	Стоимость основных производственных фондов, млн.руб.
1	2
1	16,5
2	8,8
3	22,7
4	12,7
5	3,9
6	4,6
7	4,5
8	5,9
9	9,6
10	1,9
11	5,1
12	1,6
13	4,4
14	2,8
15	3,6
16	7,6
17	4,1
18	1,4
19	17,3
20	8,9
21	4,2
22	4,6
23	1,1
24	6,2

Номер предприятия (продолжение)	Стоимость основных производственных фондов, млн.руб. (продолжение)
25	7,1
26	11,9
27	5,5
28	3,3
29	10,8
30	8,8
31	1,6
32	6
33	6,6
34	4,7
35	9,1
36	11
37	13,2
38	13,6
39	30,2
40	25,8
41	8
42	4,5
43	23,5
44	23,7
45	7,1
46	14,2
47	5,1
48	18,1
49	7,3
50	6,4
51	3,6
52	16,2

Основные расчетные формулы

1. Средняя арифметическая простая:

$$\bar{x} = \frac{\sum x_i}{n}$$

2. Средняя арифметическая взвешенная:

$$\bar{x} = \frac{\sum x_i f_i}{\sum f_i}$$

3. Мода для дискретного ряда:

- значение в ряде распределения (x_i), которому соответствует максимальная частота f_i .

4. Мода для интервального ряда:

$$x_{mod} = x_0 + h \frac{(f_m - f_{m-1})}{(f_m - f_{m-1}) + (f_m - f_{m+1})},$$

где x_0 – нижняя граница модального интервала,

h – шаг модального интервала,

f_m, f_{m-1}, f_{m+1} – частоты, соответственно: модального интервала, интервала, предшествующего модальному, и интервала, следующего за модальным.

5. Медиана для дискретного ряда:

$$x_{med} = \begin{cases} x(k), & n = 2k + 1 \\ \frac{x(k) + x(k+1)}{2}, & n = 2k \end{cases}$$

где $x(k)$ и $x(k+1)$ – порядковые статистики с порядковыми номерами соответственно k и $k+1$, т.е. значения признака в вариационном (упорядоченном) ряду распределения.

6. Медиана для интервального ряда:

$$x_{med} = x_0 + h \frac{\frac{\sum f_i}{2} - f_{m-1}^n}{f_m},$$

где x_0 – нижняя граница медианного интервала,

h – шаг медианного интервала,

f_{m-1}^n – накопленная частота интервала, находящегося перед медианным интервалом;

f_m – частота медианного интервала.

7. Нижний квартиль:

$$Q_1 = X_{Q_1} + \frac{h_{Q_1} * (\frac{1}{4} \sum f_i - f^{n_{Q_1-1}})}{f_{Q_1}}$$

X_{Q_1} – нижняя часть интервала, содержащего нижний квартиль;

h_{Q_1} – шаг интервала;

$\frac{1}{4} \sum f$ – четверть признаков всей совокупности;

$f^{n_{Q_1-1}}$ – накопленная частота интервала, предшествующего интервалу с нижним квартилем;

f_{Q_1} – частота интервала, содержащего нижний квартиль.

8. Верхний квартиль:

$$Q_3 = X_{Q_3} + \frac{h_{Q_3} * (\frac{3}{4} \sum f_i - f^{n_{Q_3-1}})}{f_{Q_3}}$$

X_{Q_3} – нижняя часть интервала содержащего верхний квартиль;

h_{Q_3} – величина интервала;

$\frac{3}{4} \sum f$ – 75 % признаков всей совокупности;

$f^{n_{Q_3-1}}$ – накопленная частота интервала, предшествующего интервалу с верхним квартилем;

f_{Q_3} – частота интервала, содержащего верхний квартиль.

9. Размах вариации (R):

$$R = x_{max} - x_{min} = x(n) - x(k),$$

где $x(l)$ и $x(n)$ – порядковые статистики с порядковыми номерами соответственно l и n , т.е. минимальное и максимальное значения признака в вариационном (упорядоченном) ряду распределения.

10. Среднее линейное отклонение (\bar{d}):

$$\bar{d} = \frac{\sum (x_i - \bar{x})}{n} \text{ - простое,}$$

$$\bar{d} = \frac{\sum (x_i - \bar{x}) f_i}{\sum f_i} \text{ - взвешенное.}$$

11. Дисперсия (S^2):

$$S^2 = \frac{\sum (x_i - \bar{x})^2}{n} \text{ - простая,}$$

$$S^2 = \frac{\sum (x_i - \bar{x})^2 f_i}{\sum f_i} \text{ - взвешенная.}$$

12. Среднеквадратическое отклонение (S):

$$S = \sqrt{S^2}$$

13. Коэффициент вариации (V_n):

$$V_n = \frac{S}{\bar{x}} * 100 \%$$