13. Отклонения

Попробуем узнать, как числа некоторого набора расположены по отношению к своему среднему арифметическому. Зная только размах, разность между наибольшим и наименьшим значением, мы не можем судить о том, как расположены числа в имеющемся наборе. Для примера возьмем набор 1, 6, 7, 9, 12. Вычислим среднее арифметическое: $(1+6+7+9+12):5=7$. Найдем отклонение каждого числа от среднего:

$$1-7=-6, \quad 6-7=-1, \quad 7-7=0, \quad 9-7=2, \quad 12-7=5.$$

Получился новый набор, который состоит из отклонений. Если число меньше среднего, то его отклонение отрицательно, если число больше среднего, то его отклонение положительно. В одном случае — для числа 7, которое совпало со средним арифметическим,— отклонение равно нулю. По набору отклонений можно судить о том, насколько разнообразны числа в наборе. Если отклонения малы, то числа в наборе расположены близко к среднему арифметическому. А если среди отклонений есть большие по модулю, то числа в наборе сильно разбросаны.

Для любого набора, если только не все числа в нем равны, часть отклонений будет положительна, а часть — отрицательна. При этом сумма всех отклонений равна 0. Убедимся в этом на нашем примере:

$$-6-1+0+2+5=0.$$

В этом состоит основное свойство отклонений: **сумма отклонений чисел от среднего арифметического этих чисел равна нулю.**

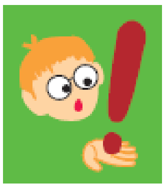
В этом пункте рассказывалось об отклонениях величины от ее среднего значения. Кроме того, мы узнали, что сумма всех отклонений в наборе от среднего равна нулю.

14. Дисперсия

Наиболее полной характеристикой разброса набора чисел является набор их отклонений от среднего арифметического. Но когда набор чисел велик, рассматривать набор отклонений практически неудобно. Нужно описать разнообразие чисел в наборе одной характеристикой, одним числом.

Размах — слишком грубая мера разброса чисел в наборе, поскольку учитывает только два из них — наименьшее и наибольшее. Можно попробовать взять «среднее отклонение». Но сумма отклонений всегда равна нулю, поэтому среднее арифметическое отклонений тоже равно нулю и его нельзя использовать как меру разброса.

Чтобы судить о разбросе, принято складывать не сами отклонения, а их квадраты. Квадраты отклонений неотрицательны, поэтому сумма квадратов отклонений зависит только от абсолютных величин отклонений, а не от их знаков. Чем больше отклонения чисел от среднего арифметического, тем больше будет сумма квадратов отклонений. Для того чтобы мера разброса чисел не зависела от их количества в наборе, в качестве такой меры берут среднее арифметическое квадратов отклонений. Эту величину называют ***дисперсией***.



Определение. Среднее арифметическое квадратов отклонений от среднего значения называется в статистике ***дисперсией*** набора чисел.

Пример 1. Снова обратимся к таблице производства пшеницы в России. Мы нашли, что среднее производство пшеницы за

период 1995—2001 гг. составило 35,5 млн. тонн в год. Вычислим дисперсию. Составим таблицу, разместив данные по производству не в строке, а в столбце. Вычислим отклонения от среднего и их квадраты. Полученные числа занесем в два новых столбца.

Таблица 8. Производство пшеницы в России в 1995—2001 гг., млн. тонн

Год	Производство	Отклонение от среднего	Квадрат отклонения
1995	30,1	−5,4	29,16
1996	34,9	−0,6	0,36
1997	44,3	8,8	77,44
1998	27,0	−8,5	72,25
1999	31,0	−4,5	20,25
2000	34,5	−1,0	1,00
2001	47,0	11,5	132,25

Для расчета дисперсии следует сложить все значения в столбце «Квадрат отклонения» и разделить на количество слагаемых:

$$(29{,}16+0{,}36+77{,}44+72{,}25+20{,}25+1{,}00+132{,}25):7=47{,}53.$$

Пример 2. Покажем на простом примере, как дисперсия характеризует разброс наблюдений. Возьмем два набора чисел 1, 2, 3 и 0, 2, 4. Среднее арифметическое значение обоих наборов равно 2. Для обоих наборов вычислим отклонения и квадраты отклонений и все данные занесем в таблицу 9.

Таблица 9

1-й набор	Отклонение от среднего	Квадрат отклонения	2-й набор	Отклонение от среднего	Квадрат отклонения
1	−1	1	0	−2	4
2	0	0	2	0	0
3	1	1	4	2	4

Дисперсия первого набора: $(1+0+1):3=\frac{2}{3}$.

Дисперсия второго набора: $(4+0+4):3=2\frac{2}{3}$.

Числа в первом наборе расположены более кучно — ближе друг к другу и к своему среднему, — чем числа во втором наборе. Поэтому дисперсия первого набора получилась меньше, чем второго.

Пример 3. Континентальный климат отличается от умеренного более резкими изменениями температуры в течение года. В районах с континентальным климатом жаркое лето и очень холодная зима. С помощью дисперсии различия между двумя видами климата можно выразить количественно. Сравним для примера изменение температур в течение года в Москве и Киеве, где климат умеренный, с изменением температур в Новосибирске и Хабаровске, где климат континентальный. В таблице 10 приведены средние месячные температуры за 80 лет в Москве, Киеве, Новосибирске и Хабаровске.

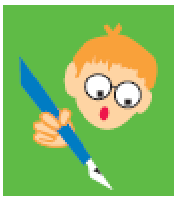
Таблица 10. Средние месячные температуры, °C

Месяцы	Москва	Киев	Новосибирск	Хабаровск
1	−9,3	−5,9	−19,0	−22,3
2	−8,6	−5,2	−17,2	−17,2
3	−3,4	−0,4	−10,7	−8,5
4	5,1	7,5	−0,1	3,1
5	12,4	14,7	10,0	11,1
6	16,7	17,8	16,3	17,4
7	18,4	19,8	18,7	21,1
8	16,6	18,7	16,0	20,0
9	10,9	13,9	9,9	13,9
10	4,4	7,5	1,5	4,7
11	−2,0	1,2	−9,7	−8,1
12	−6,8	−3,5	−16,9	−18,5
Среднее за год	4,5	6,0	−0,1	−1,4
Дисперсия	98,9	86,5	185,2	228,8

Дисперсии этих четырех рядов чисел различны. Для Москвы и Киева это 98,9 и 86,5, для Новосибирска и Хабаровска это 185,2 и 228,8.

Вы видите, что дисперсии для рядов месячных температур в умеренном и континентальном климате значительно различаются.

Мы познакомились еще с одним показателем рассеивания набора — с дисперсией.



Упражнения

1. Для данных чисел вычислите среднее значение. Составьте таблицу отклонений от среднего и квадратов отклонений от среднего и вычислите дисперсию:

а) −1, 0, 4; б) 2, 3, 7; в) −3, 1, 2, 4; г) 2, 6, 7, 5;
д) −2, −1, 1, 2, 5; е) −1, −3, −2, 3, 3.

2. Даны два набора чисел. Отметьте их на числовой прямой. Вычислите дисперсию каждого из этих наборов. Дисперсия какого набора больше?

а) 2, 3, 7 и 1, 2, 3; б) 2, 3, 4, 7 и 1, 5, 6, 8.

3. Даны два набора чисел. Отметьте их на числовой прямой. Вычислите дисперсию каждого из этих наборов. Сравните дисперсии:

а) 2, 3, 4 и 6, 7, 8; б) 3, 5, 7, 9 и 12, 14, 16, 18.

Домашнее задание: Выписать определения и выводы в тетрадь. Выполнить упражнения.