

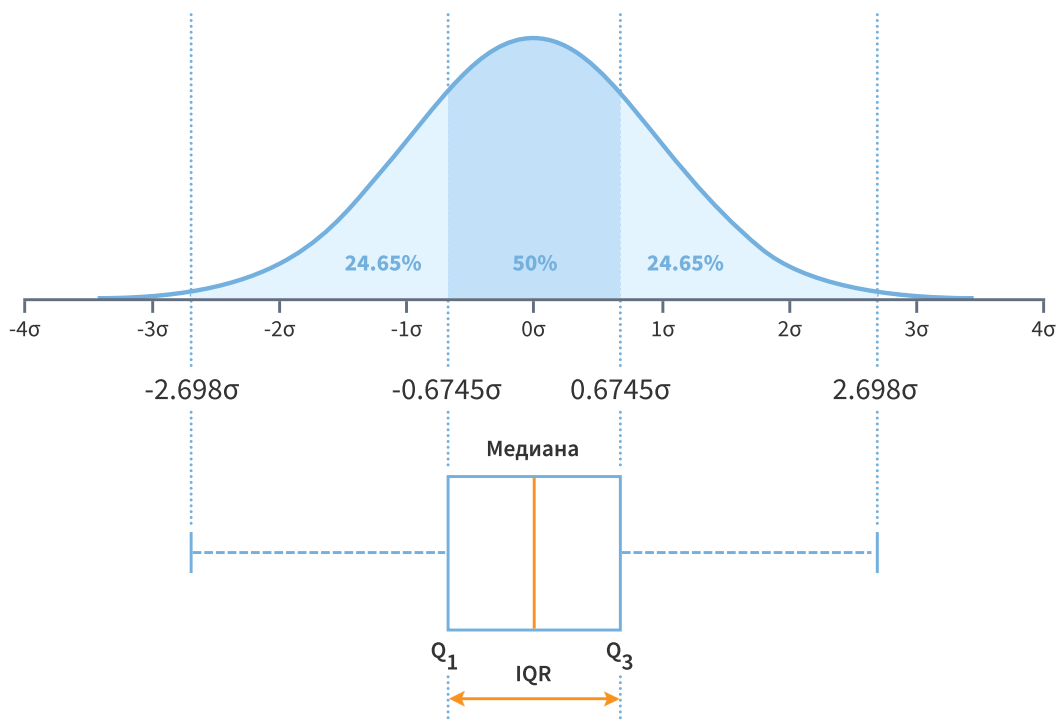
Интерквартильный размах (Interquartile range)

Синонимы: IQR, Midspread, Middle 50%, Fourth spread, H-spread

Разделы: [Метрики](#)

Интерквартильный размах представляет собой порядковую статистику, численно равную разности между 1-м и 3-м [квантилями](#) распределения или 25-м и 75-м процентилями. Можно так же сказать, что интерквартильный размах это половина [выборки](#), центрированная относительно [медианы](#).

Интерквартильный размах является удобным показателем изменчивости [признака](#) для асимметричных распределений или наборов данных с [аномальными значениями](#). Иными словами его можно рассматривать в качестве аналога [дисперсии](#), который является [робастным](#) к выбросам в данных.



На рисунке интерквартильный размах проиллюстрирован с помощью диаграммы [ящик с усами](#). На нем буквами Q_1 и Q_3 обозначены 1-й и 3-й квантили соответственно. Ширина «ящика» показывает интерквартильный размах.

Рассмотрим пример вычисления интерквартильного размаха.

№	Исходные данные	Ранжированные данные
1	75	7
2	31	7
3	7	31
4	87	32
5	115	47
6	121	75
7	47	87
8	155	115
9	116	116
10	119	119
11	177	121
12	7	155
13	32	177

Исходные значения представлены во второй колонке таблицы (13 значений). Чтобы вычислить порядковые статистики, нужно ранжировать значения по возрастанию или убыванию. Сначала найдем медиану — значение ранжированной последовательности справа и слева от которого будет одинаковое число элементов. Несложно увидеть, что в нашем примере $Me = 87$.

Затем в каждом интервале слева и справа от медианы нужно найти значения, которые делят эти интервалы на две равные части. Это и будут 2-й и 3-й квартили соответственно. Поскольку в каждом из интервалов оказалось четное количество значений (поскольку общее число значений ряда нечетное), в качестве квартилей следует взять два срединных значения и усреднить их.

Таким образом, в нашем примере первый квартиль будет $Q_1 = (31 + 32)/2 = 31.5$, а третий $Q_3 = (119 + 121)/2 = 120$. Следовательно, интерквартильный размах будет $IRQ = Q_3 - Q_1 = 88.5$.

В анализе данных интерквартильный размах широко используется при оценке качества данных с целью обнаружения в них выбросов и аномальных значений. Например, может быть определено условие, что все значения набора данных, которые отклоняются от медианы больше, чем на $3 IRQ$, считаются выбросами и к ним должны применяться соответствующие методы обработки.

В этом случае интерквартильный размах может быть использован для обнаружения аномальных значений как альтернатива среднеквадратическому отклонению, которое целесообразно применять только для нормально распределенных данных. В то же время интерквартильный размах является непараметрической оценкой, поэтому вообще не требует предположений относительно распределения и может применяться для любых данных.

В аналитической платформе Loginom есть специализированный обработчик Редактирование выбросов, в котором интерквартильный размах может быть вычислен и использован для обнаружения выбросов в данных.