

Know when to discard an outlier data point in HPLC by using the Q Test - How To

For example, if you prepare six replicate samples and obtain the following peak areas for your **analyte**:

106.5
 104.2
 103.7
 107.1
 99.2
 104.7

The values are fairly close together, but what about 99.2? Should this data point be rejected or kept? i.e. Is it a legitimate data point or an outlier? To answer this, we use a test to determine — at a defined confidence limit — the reliability of the data point, known as the **Q Test**.

We need to compute two values, $Q_{\text{calculated}}$ and Q_{table} . $Q_{\text{calculated}}$ is obtained as follows:

$$Q_{\text{calculated}} = \text{gap} / \text{range}$$

...where *gap* is the absolute difference between the suspect data point and its nearest neighbor and *range* is the difference between the highest and lowest values in the data set.

$$\begin{aligned} \text{gap} &= 103.7 - 99.2 = 4.5 \\ \text{range} &= 107.1 - 99.2 = 7.9 \\ Q_{\text{calculated}} &= 4.5 / 7.9 = 0.57 \end{aligned}$$

To find Q_{table} , we look it up in the following table:

Number of values:	3	4	5	6	7	8	9	10
Q _{90%} :	0.941	0.77	0.64	0.56	0.507	0.468	0.437	0.412
Q _{95%} :	0.97	0.83	0.71	0.625	0.568	0.526	0.493	0.466
Q _{99%} :	0.994	0.93	0.82	0.74	0.68	0.634	0.598	0.568

We have 6 data points, so at the 95% confidence level, $Q_{table} = 0.625$. The criteria for acceptance or rejection are as follows:

If $Q_{calculated} < Q_{table}$, accept the data point

If $Q_{calculated} > Q_{table}$, reject the data point

Since $0.57 < 0.625$, the data point can be kept with confidence.