

# **Peirce's criterion for the elimination of suspect experimental data**

Stephen M. Ross, Ph.D.

*Journal of Engineering Technology*, Fall 2003.

Professor of Mechanical Engineering, University of New Haven, 300 Orange Ave., West Haven, CT 06516

## **ABSTRACT**

Peirce's criterion is a rigorous method based on probability theory that can be used to eliminate data "outliers" or spurious data in a rational way. Currently, another method called Chauvenet's criterion is used in many educational institutions and laboratories to perform this function. Although Chauvenet's criterion is well established, it makes an arbitrary assumption concerning the rejection of the data. Peirce's criterion does not make this arbitrary assumption. In addition, Chauvenet's criterion makes no distinction between the case of one or several suspicious data values whereas Peirce's criterion is a rigorous theory that can be easily applied in the case of *several suspicious data values*. In this paper, an example is given showing that Peirce's and Chauvenet's criterion give different results for the particular set of data presented.

## **INTRODUCTION**

Peirce's criterion has been buried in the scientific literature for approximately 150 years. It is virtually unknown today in the scientific community. In its place, Chauvenet's criterion is commonly used for rational elimination of "outlier" data by government laboratories, (e.g., Environmental Protection Agency, U.S. Army Corps of Engineers,

Agency for Toxic Substances and Disease Registry, Institute for Telecommunication Sciences), industry (e.g., Boeing, Sikorsky), foreign laboratories (e.g., Laboratoire National Henri Becquerel, Joint Astronomy Centre), and universities (e.g., research and courses at University of Michigan, Texas A&M, University of California, Vanderbilt, University of Alberta, Ohio State).

Methods of elimination of data “outliers” are useful for anyone working in industry or in an educational institution where statistical information concerning product runs or experimental data is of interest. In an engineering, technology or science program, laboratory courses in chemistry, physics and engineering can, and do, find use for rational spurious data elimination.

In the BSME program at the University of New Haven, we have used Chauvenet's criterion in our instrumentation and fluid/thermal laboratory courses for many years. Other universities have similarly used this criterion in their undergraduate laboratories. Typically, students take several measurements of a quantity, say pressure, at one setting (meaning the experimental conditions are maintained at the same level). Assuming the systematic errors are negligible, each measurement will vary slightly due to random errors (e.g., reading instrument values, flow rate may change slightly, etc.). Often, however, one or two datum points seem to fall “far” outside the range of the others obtained. These outliers greatly impact the mean and standard deviation of the measurements. A data elimination method can be used to obtain a realistic average value of pressure and an “uncertainty” in the true value given by the standard deviation.

## BACKGROUND

In the course of taking experimental data, there is often the situation in which one or more of the data values seems intuitively to be well outside of the norm expected.

Whatever the reasons for this occurrence, whether due to a temporary equipment glitch or not, a rational method can be used to eliminate the data point(s) in question.

This problem was originally addressed by Peirce <sup>1</sup>, Gould <sup>2</sup>, and later by Chauvenet <sup>3</sup> in 1863. All of them used elements of probability theory to develop a logical method to eliminate suspicious data from a set of measurements.

Chauvenet's criterion is in common use today for elimination of suspect data. Among other texts, it can be found in different, but equivalent, form in Holman's book on experimental methods <sup>4</sup> and Taylor's text <sup>5</sup> on error analysis. In Chauvenet's method, there is an arbitrary assumption that a measurement may be rejected if the probability of obtaining the deviation from the mean for that value is less than the inverse of twice the number of measurements. The theoretical development of Peirce's criterion does not make any such assumption. In addition, whereas Chauvenet's criterion makes no distinction between the case of one or several suspicious data values, Peirce's criterion is a rigorous theory that can be easily applied in the case of *several suspicious data values* using the table in this paper.

Chauvenet himself believed that Peirce's work was more rigorous and could be applied more generally, and in Chauvenet's words, "*For the general case..... when there are several unknown quantities and several doubtful observations, the modifications which the rule (meaning his own criterion) requires renders it more troublesome than Peirce's*

*formula.....What I have given may serve the purpose of giving the reader greater confidence in the correctness and value of Peirce's Criterion".* <sup>3</sup>

Peirce's criterion is an in-depth derivation formulated directly from the theory of probability. It was based on the following principle: "*...the proposed observations should be rejected when the probability of the system of errors (my note: the actual deviations from the mean) obtained by retaining them is less than that of the system of errors obtained by their rejection multiplied by the probability of making so many, and no more, abnormal observations*". <sup>1</sup> The actual method of calculation used by Peirce was mathematically cumbersome to use and a later paper by Gould took Peirce's criterion and presented it in a more easily usable format with tables he derived from Peirce's work.

The table of Peirce's criterion for one measured quantity is reproduced here in modified form covering up to a maximum of 60 total measurements. It is denoted as "Peirce's Criterion Table" in this paper. <sup>6</sup>

In the table, R represents the ratio of the maximum allowable deviation of a measured value from the data mean to the standard deviation. That is,

$$R = \frac{|x_i - x_m|_{\max}}{\sigma} \quad (1)$$

where  $\sigma$  is the "sample" standard deviation of the complete data set,

$x_i$  is a measured data value, and

$x_m$  is the mean of the data set.

## PEIRCE'S METHOD

The method to determine whether data values should be rejected involves the procedure outlined below:

1. Calculate the mean and the sample standard deviation of the complete data set.
2. Obtain  $R$  corresponding to the number of measurements taken from Peirce's table.

Assume the case of one doubtful observation first, even if there appears to be more than one.

3. Calculate the maximum allowable deviation:  $|x_i - x_m|_{max}$ .
4. For any suspicious data measurements, obtain  $|x_i - x_m|$ .
5. Eliminate the suspicious measurements if:

$$|x_i - x_m| > |x_i - x_m|_{max} \quad (2)$$

6. If this results in the rejection of one measurement, assume the case of two doubtful observations, keeping the original values of the mean and standard deviation, and the original number of measurements. Go to step 8.
7. If more than one measurement is rejected in the above test, assume the next highest value of doubtful observations. For example, if two measurements are rejected in

step 5, assume the case of three doubtful observations, keeping the original values of the mean, standard deviation, and the original number of measurements as the process is continued.

8. Repeat the above calculations (steps 2 – 5), sequentially increasing the number of doubtful measurement possibilities, until no more data measurements need to be eliminated.
9. Now obtain the new value of the mean and sample standard deviation of the reduced data set.

### **EXAMPLE: PEIRCE'S CRITERION**

In this example, one of the datum points will be rejected using Peirce's criterion. A comparison using Chauvenet's criterion gives a different result for this particular set of data. None of the ten data values are eliminated using Chauvenet's criterion.

Ten pressure measurements are taken in an experiment at one setting. The gage values are 101.2, 90.0, 99.0, 102.0, 103.0, 100.2, 89.0, 98.1, 101.5, 102.0 kPa . The pressures of 89.0 and 90.0 appear suspect. We follow Peirce's method in the steps below.

1. Determine the sample standard deviation and mean for the complete set:  $\sigma = 5.02$ ,  
 $x_m = 98.6$ .
2. Obtain  $R$  from the table for one measured quantity assuming one doubtful observation and ten measurements:  $R = 1.878$ . (Note that it is possible that more than one measurement may be eliminated in this first round of checking.)
3. Calculate the maximum allowable deviation:

$$|x_i - x_m|_{\max} = \sigma R = 5.02 (1.878) = 9.43.$$

4. Obtain the actual deviations for the suspicious measurements:

$$|89 - x_m| = |89.0 - 98.6| = 9.60.$$

$$|90 - x_m| = |90.0 - 98.6| = 8.60.$$

5. Check for elimination of the suspicious data using equation 2:

$$9.60 > 9.43 \quad \text{Eliminate this datum measurement (89.0).}$$

$$8.60 < 9.43 \quad \text{Retain this datum measurement (90.0).}$$

6. Now repeat the above assuming two doubtful data measurements.<sup>7</sup> From the table, second column, still using the ten observation row:  $R = 1.570$ . Obtain the maximum allowable deviation using the original standard deviation and mean:  $|x_i - x_m|_{\max} = \sigma R = 5.02 (1.570) = 7.88$ .

$$\text{Check for elimination of suspect point 90.0: } |90 - x_m| = |90.0 - 98.6| = 8.60.$$

Since  $8.60 > 7.88$ , eliminate the measurement 90.0. No other measurements are eliminated at this step.

7. To consider rejecting a third measurement, repeat this process using the third column from the table with the row still corresponding to ten observations:  $R = 1.380$ . Obtain the maximum allowable deviation using the original standard deviation and mean:  $|x_i - x_m|_{\max} = \sigma R = 5.02 (1.380) = 6.93$ .

A check of  $|x_i - x_m|$  for the remaining eight observations shows that no other measurements need to be eliminated. *This concludes the process of elimination.*

8. For the final result, re-calculate the sample standard deviation and the mean for the reduced data set of 8 remaining observations (i.e., the original set with points 89.0 and 90.0 eliminated):  $\sigma = 1.66$ ,  $x_m = 100.9$ .

An analysis using Chauvenet's criterion for the same data set does not reject the two suspect data values above. From table 3.5, P. 79 from Holman's text<sup>5</sup>, we have the maximum acceptable deviation to standard deviation ratio of 1.96. Thus,

$|x_i - x_m|_{max} = 5.02 (1.96) = 9.84$ . Referring back to step 4 above, we have

$|89 - x_m|$  and  $|90 - x_m|$  equal to 9.60 and 8.60, both of which are acceptable; none of the ten datum values are eliminated.

## **SUMMARY AND CONCLUSION**

Chauvenet's criterion, commonly used today for the rational elimination of datum "outliers" was preceded by the more rigorous, and more general, Peirce's criterion.

It is recommended that Peirce's criterion be used instead of Chauvenet's criterion for the elimination of suspect data for the following reasons:

- Peirce 's formulation is more rigorous.
- Peirce's criterion does not make an arbitrary assumption concerning the rejection of data.
- Peirce's criterion theoretically accounts for the case where there is more than one suspect observation. Chauvenet's criterion does not.
- Peirce's method can be easily followed using the table in this paper.

A modified version of Peirce's criterion in table form along with a detailed example illustrating how to apply the criterion for a set of data measurements is reproduced in this paper. Peirce's criterion gives a different result than Chauvenet's criterion for this



particular set of data. The table contains values for up to 60 total measurements with one to nine doubtful observations for one type of measured unknown quantity.

The author wishes to thank Professors John Sarris and John R. Taylor for their helpful criticism in preparation of this paper.

## FOOTNOTES AND REFERENCES

- <sup>1</sup> Benjamin Peirce, " Criterion for the rejection of doubtful observations, " *Astronomical Journal* II, **45**, 161- 163 (1852).
- <sup>2</sup> B.A. Gould, " On Peirce's criterion for the rejection of doubtful observations, with tables for facilitating its application, " *Astronomical Journal* IV, **83**, 81 - 87 (1855).
- <sup>3</sup> William Chauvenet, *A Manual of Spherical and Practical Astronomy V.II* , Lippincott, Philadelphia, 1<sup>st</sup> Ed (1863); Reprint of 1891 5<sup>th</sup> Ed: Dover, NY (1960).
- <sup>4</sup> John Taylor, *Error Analysis* , 2<sup>nd</sup> Ed, pp. 166 – 170, University Sci. Books, Sausalito, CA (1997).
- <sup>5</sup> J.P. Holman, *Experimental Methods for Engineers*, 7<sup>th</sup> Ed, pp. 78 – 80, McGraw Hill (2001).
- <sup>6</sup> The case of measurements for two different quantities appears to be relevant only to certain types of astronomical data. The original, and only, example that pertains to this case appears in Peirce's paper. The true meaning of the term "two unknown quantities" is not clear; however, it seems to pertain to both direct and indirect angle measurements. This same unclear example is exactly repeated in Gould's paper and Chauvenet's book. For typical experimental measurements in various disciplines, only the table for one unknown quantity is useful.
- <sup>7</sup> If both measurements were eliminated in step 5, the process would move on to the third column of the table in step 6 of the example.

**PEIRCE'S CRITERION TABLE****ONE MEASURED QUANTITY****VALUES OF R**

Total number

of observations

Number of doubtful observations

	<b>1</b>	<b>2</b>	<b>3</b>	<b>4</b>	<b>5</b>	<b>6</b>	<b>7</b>	<b>8</b>	<b>9</b>
<b>3</b>	1.196								
<b>4</b>	1.383	1.078							
<b>5</b>	1.509	1.200							
<b>6</b>	1.610	1.299	1.099						
<b>7</b>	1.693	1.382	1.187	1.022					
<b>8</b>	1.763	1.453	1.261	1.109					
<b>9</b>	1.824	1.515	1.324	1.178	1.045				
<b>10</b>	1.878	1.570	1.380	1.237	1.114				
<b>11</b>	1.925	1.619	1.430	1.289	1.172	1.059			
<b>12</b>	1.969	1.663	1.475	1.336	1.221	1.118	1.009		
<b>13</b>	2.007	1.704	1.516	1.379	1.266	1.167	1.070		
<b>14</b>	2.043	1.741	1.554	1.417	1.307	1.210	1.120	1.026	
<b>15</b>	2.076	1.775	1.589	1.453	1.344	1.249	1.164	1.078	
<b>16</b>	2.106	1.807	1.622	1.486	1.378	1.285	1.202	1.122	1.039
<b>17</b>	2.134	1.836	1.652	1.517	1.409	1.318	1.237	1.161	1.084
<b>18</b>	2.161	1.864	1.680	1.546	1.438	1.348	1.268	1.195	1.123
<b>19</b>	2.185	1.890	1.707	1.573	1.466	1.377	1.298	1.226	1.158
<b>20</b>	2.209	1.914	1.732	1.599	1.492	1.404	1.326	1.255	1.190

<b>21</b>	2.230	1.938	1.756	1.623	1.517	1.429	1.352	1.282	1.218
<b>22</b>	2.251	1.960	1.779	1.646	1.540	1.452	1.376	1.308	1.245
<b>23</b>	2.271	1.981	1.800	1.668	1.563	1.475	1.399	1.332	1.270
<b>24</b>	2.290	2.000	1.821	1.689	1.584	1.497	1.421	1.354	1.293
<b>25</b>	2.307	2.019	1.840	1.709	1.604	1.517	1.442	1.375	1.315
<b>26</b>	2.324	2.037	1.859	1.728	1.624	1.537	1.462	1.396	1.336
<b>27</b>	2.341	2.055	1.877	1.746	1.642	1.556	1.481	1.415	1.356
<b>28</b>	2.356	2.071	1.894	1.764	1.660	1.574	1.500	1.434	1.375
<b>29</b>	2.371	2.088	1.911	1.781	1.677	1.591	1.517	1.452	1.393
<b>30</b>	2.385	2.103	1.927	1.797	1.694	1.608	1.534	1.469	1.411
<b>31</b>	2.399	2.118	1.942	1.812	1.710	1.624	1.550	1.486	1.428
<b>32</b>	2.412	2.132	1.957	1.828	1.725	1.640	1.567	1.502	1.444
<b>33</b>	2.425	2.146	1.971	1.842	1.740	1.655	1.582	1.517	1.459
<b>34</b>	2.438	2.159	1.985	1.856	1.754	1.669	1.597	1.532	1.475
<b>35</b>	2.450	2.172	1.998	1.870	1.768	1.683	1.611	1.547	1.489
<b>36</b>	2.461	2.184	2.011	1.883	1.782	1.697	1.624	1.561	1.504
<b>37</b>	2.472	2.196	2.024	1.896	1.795	1.711	1.638	1.574	1.517
<b>38</b>	2.483	2.208	2.036	1.909	1.807	1.723	1.651	1.587	1.531
<b>39</b>	2.494	2.219	2.047	1.921	1.820	1.736	1.664	1.600	1.544
<b>40</b>	2.504	2.230	2.059	1.932	1.832	1.748	1.676	1.613	1.556
<b>41</b>	2.514	2.241	2.070	1.944	1.843	1.760	1.688	1.625	1.568
<b>42</b>	2.524	2.251	2.081	1.955	1.855	1.771	1.699	1.636	1.580
<b>43</b>	2.533	2.261	2.092	1.966	1.866	1.783	1.711	1.648	1.592
<b>44</b>	2.542	2.271	2.102	1.976	1.876	1.794	1.722	1.659	1.603
<b>45</b>	2.551	2.281	2.112	1.987	1.887	1.804	1.733	1.670	1.614
<b>46</b>	2.560	2.290	2.122	1.997	1.897	1.815	1.743	1.681	1.625
<b>47</b>	2.568	2.299	2.131	2.006	1.907	1.825	1.754	1.691	1.636

<b>48</b>	2.577	2.308	2.140	2.016	1.917	1.835	1.764	1.701	1.646
<b>49</b>	2.585	2.317	2.149	2.026	1.927	1.844	1.773	1.711	1.656
<b>50</b>	2.592	2.326	2.158	2.035	1.936	1.854	1.783	1.721	1.666
<b>51</b>	2.600	2.334	2.167	2.044	1.945	1.863	1.792	1.730	1.675
<b>52</b>	2.608	2.342	2.175	2.052	1.954	1.872	1.802	1.740	1.685
<b>53</b>	2.615	2.350	2.184	2.061	1.963	1.881	1.811	1.749	1.694
<b>54</b>	2.622	2.358	2.192	2.069	1.972	1.890	1.820	1.758	1.703
<b>55</b>	2.629	2.365	2.200	2.077	1.980	1.898	1.828	1.767	1.711
<b>56</b>	2.636	2.373	2.207	2.085	1.988	1.907	1.837	1.775	1.720
<b>57</b>	2.643	2.380	2.215	2.093	1.996	1.915	1.845	1.784	1.729
<b>58</b>	2.650	2.387	2.223	2.101	2.004	1.923	1.853	1.792	1.737
<b>59</b>	2.656	2.394	2.230	2.109	2.012	1.931	1.861	1.800	1.745
<b>60</b>	2.663	2.401	2.237	2.116	2.019	1.939	1.869	1.808	1.753

We then used manually classified data sets of repositories to train classifiers capable of predicting if a given GitHub repository contains an engineered software project. The performance of the classifiers was evaluated using a set of 200 repositories with known ground truth classification. We also compared the performance of the classifiers to other approaches to classification (e.g. number of GitHub Stargazers) and found our classifiers to outperform existing approaches. We found stargazers-based classifier (with 10 as the threshold for number of stargazers) to exhibit high precision (97%) ...

1. Ross SM (2003) Peirce's criterion for the elimination of suspect experimental data. *J Eng Technol* 20(2):38-41. Google Scholar. Peirce's criterion is a method, devised by Benjamin Peirce, that may be used to eliminate suspect experimental data using probability theory. Description. For scientists, engineers and others involved in real data collection, the situation often arises in which one or more of the measured values appears to be outside the usual range. The temptation to ignore this data with rationalisations such as blaming faulty recording equipment ( the equipment had a power surge, there was dirt in the lens) should be resisted. Instead of arbitrarily dropping data, Peirce's Criterion (1) may be applied.

2. Stephen Ross, "Peirce's Criterion for the Elimination of Suspect Experimental Data", *J. Engr. Technology*, Fall, 2003. [<http://newton.newhaven.edu/sross/piercescriterion.pdf>].

3. D.M. Hawkins (1980). B. Peirce, "Criterion for the rejection of doubtful observations", *Astronomical Journal* II, 45, 161-163, 1852

12. C. Reimann, P. Filzmoser and R.G. Garrett, "Background and threshold: critical comparison of. methods of determination", *Science of the Total Environment*, 346, 1-16, 2005

13. S.M. Ross, "Peirce's criterion for the elimination of suspect experimental data", *Journal of. Engineering technology*, Fall, 2003

14. P.J. Rousseeuw and A.M. Leroy, *Robust regression and outlier detection*, John Wiley & Son