**Rejection of Measurements**

Often when repeating measurements, one value appears to be spurious and we would like to throw it out. Also, when taking a series of measurements, sometimes one value appears "out of line". Here we discuss some guidelines on rejection of measurements.

It is important to emphasize that the whole topic of rejection of measurements is awkward. Some scientists feel that the rejection of data is never justified unless there is external evidence that the data in question is incorrect. Other scientists attempt to deal with this topic by using quasi-objective rules such as [Chauvenet's Criterion](http://instructor.physics.lsa.umich.edu/ip-labs/tutorials/errors/illerr.html%22%20%5Ct%20%22_ea). Still others, often incorrectly, throw out any data that appear to be incorrect. In this section, some principles and guidelines are presented.

First, we note that it is incorrect to expect each and every measurement to overlap within errors. For example, if the error in a particular quantity is characterized by the standard deviation, we only expect 68% of the measurements from a normally distributed population to be within one standard deviation of the mean. Ninety-five percent of the measurements will be within two standard deviations, 99% within three standard deviations, etc., but we never expect 100% of the measurements to overlap within any finite-sized error for a truly Gaussian distribution.

Of course, for most experiments the assumption of a Gaussian distribution is only an approximation.

If the error in each measurement is taken to be the reading error, again we only expect most, not all, of the measurements to overlap within errors. In this case the meaning of "most", however, is vague and depends on the optimism/conservatism of the experimenter who assigned the error.

Thus, it is always dangerous to throw out a measurement. Maybe we are unlucky enough to make a valid measurement that lies ten standard deviations from the population mean. A valid measurement from the tails of the underlying distribution should not be thrown out. It is even more dangerous to throw out a suspect point indicative of an underlying physical process. Very little science would be known today if the experimenter always threw out measurements that didn't match preconceived expectations!

In general, there are two different types of experimental data taken in a laboratory and the question of rejecting measurements is handled in slightly different ways for each. The two types of data are the following:

1. A series of measurements taken with one or more variables changed for each data point. An example is the calibration of a thermocouple, in which the output voltage is measured when the thermocouple is at a number of different temperatures.
2. Repeated measurements of the same physical quantity, with all variables held as constant as experimentally possible. An example is measuring the time for a pendulum to undergo 20 oscillations and repeating the measurement five times, as in Question 10.1.

For a series of measurements (case 1), when one of the data points is out of line the natural tendency is to throw it out. But, as already mentioned, this means you are assuming the result you are attempting to measure. As a rule of thumb, unless there is a physical explanation of why the suspect value is spurious and it is no more than three standard deviations away from the expected value, it should probably be kept. Thus, throwing out a measurement is usually justified only if both of the following are true:

* A definite physical explanation of why the suspect value is spurious is found.
* The suspect value is more than three error bars away from the expected (interpolation/extrapolated) value, and no other measurement appears spurious.

For this case of a series of measurements, there is a whole range of *robust* fitting techniques which attempt to objectively determine to what degree a single datum may be ignored.

For repeated measurements (case 2), the situation is a little different. Say you are measuring the time for a pendulum to undergo 20 oscillations and you repeat the measurement five times. Assume that four of these trials are within 0.1 seconds of each other, but the fifth trial differs from these by 1.4 seconds (i.e., more than three standard deviations away from the mean of the "good" values). There is no known reason why that one measurement differs from all the others. Nonetheless, you may be justified in throwing it out. Say that, unknown to you, just as that measurement was being taken, a gravity wave swept through your region of space time. However, if you are trying to measure the period of the pendulum when there are no gravity waves affecting the measurement, then throwing out that one result is reasonable. (Although trying to repeat the measurement to find the existence of gravity waves will certainly be more fun!) So whatever the reason for a suspect value, the rule of thumb is that it may be thrown out provided that fact is well documented and that the measurement is repeated a number of times more to convince the experimenter that he/she is not throwing out an important piece of data indicating a new physical process.