# Number of Trials Needed for an Experiment: A Statistical Approach

Poovizhi M[1] and V. Madhurima[1]

[1]Department of Physics, Central University of Tamil Nadu, Thiruvarur, India.
poovizhimaruthi@gmail.com; madhurima@cutn.ac.in

*Submitted on 06-05-2026*
*Accepted on 02-06-2026*

## Abstract

Repeated measurements are fundamental to scientific methodology. However, many are unaware of the number of trials required for any given experiment. We ask and answer an important question - ”are the required number of trials fixed for all the experiments?”. We present a simple statistical framework taking the examples of length measurement using a vernier callipers and of percolation of water.

## 1 Introduction

Repeatability and reproducibility, i.e., obtaining the same results when the experiments are conducted under the same conditions, either by the same person or someone else, are crucial for scientific rigour. The natural question that follows is ”how many times?”. Students often have similar responses, and explanations, to the question of how many times they should repeat the measurement of length using vernier callipers. It is fairly common to hear 2 (check both values are same), 3 (three points make a straight line), 5 (teacher told us) or 6 (just feel like it). When the question extends to any other experiment, the common responses are the same as with Vernier Callipers, or that they do not have an idea!.

The answer to this question is not intuitive, it is deeply seated in statistics. When an experiment is performed we usually repeat the experiment before presenting the result. But have we asked ”why do we repeat that?” or ”how many times do we repeat it and is the repetition same for all the experiments”. The answer to these questions can be found from the ”law of large numbers” in statistics, which states that as the number of observations increases, the sample mean gets closer to the true population mean. In other words, with a large enough number of measurements, the average is a reliable estimate of the true value.

The number of observations required

depends on the variability of the data. If the variance is small, the sample mean stabilizes quickly and fewer observations are needed. If the variance is large, more observations are required for the sample mean to closely approximate the true mean. This idea governs the number of trials required. Let us explore this a little further.

## 2 Statistical Description of Measurements

When a physical quantity is measured repeatedly in an experiment, we usually do not obtain exactly the same value each time. Therefore, instead of relying on a single measurement, we perform the experiment multiple times and analyze the set of observations statistically.

For a set of ($n$) measurements ( $x_i$), the arithmetic mean is defined as the sum of all data divided by the number of data and is given by

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i. \tag{1}$$

The mean represents the best estimate of the true value based on the available measurements, as shown in Figure 1.

The standard deviation, which shows the spread of data around the mean, is given by

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2}. \tag{2}$$

It quantifies how much the individual measurements fluctuate about the mean and therefore reflects the consistency (precision) of the experiment,

To compare variability across different experiments, especially when the measured quantities or their magnitudes are different, the coefficient of variation is defined as

$$CV = \frac{\sigma}{\bar{x}}. \tag{3}$$

Since it expresses the spread relative to the mean, it allows meaningful comparison of precision between experiments with different scales or units. For example, in the vernier calliper measurements, the same standard deviation (e.g., 0.1 mm) represents a much larger relative uncertainty for a 10 mm object than for a 100 mm object. The coefficient of variation accounts for this by expressing the spread relative to the mean, allowing meaningful comparison of precision across measurements of different magnitudes.

The standard error is defined as the uncertainty associated with the mean and is given by

$$SE = \frac{\sigma}{\sqrt{n}}. \tag{4}$$

It indicates how precisely the mean has been determined and decreases as the number of measurements increases.

For measurements with small fluctuations, repeated readings tend to cluster around a central value and may be approximated by a normal distribution.

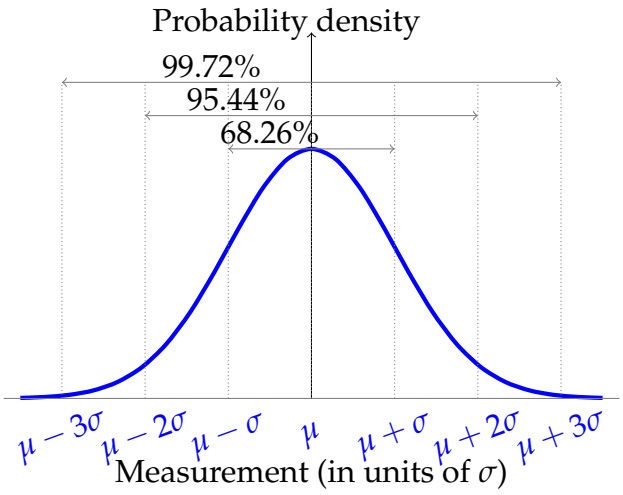


Figure 1: Normal distribution curve

## 3 Estimation of Required Number of Trials

Let us consider the case where we have obtained a large dataset. Many natural and experimental measurements approximately fit a normal distribution, which is a symmetric, bell-shaped distribution in which most measurements cluster around the mean and fewer occur as we move away from it.

For repeated measurements of a length using a vernier caliper, if the readings follow a normal distribution, the 95% confidence interval for the true length is

$$[\bar{x} \pm z \cdot SE]$$

Here, $(\bar{x})$is the average of the vernier readings and (SE) reflects how precisely that average has been determined from repeated measurements. The $(z) - value$ (1.96 for 95% confidence interval) comes from the normal distribution and represents how many standard deviations around the mean capture most of the repeated vernier measurements.

For such data, a 95% confidence interval for the mean is given by

$$\bar{x} \pm z \cdot SE. \tag{5}$$

where the z-score comes from the normal distribution table and is given by how many $x\,\sigma$ a data point is away from the mean.

| Z-score | Left tail % | Confidence level |
|---|---|---|
| 1.00 | 84.1% | 68% |
| **1.96** | **97.5%** | **95%** |
| 2.00 | 97.7% | 95.4% |
| 3.00 | 99.9% | 99.7% |

Table 1: Z value and the confidence level

For a 95% confidence interval,

$$\bar{x} \pm 1.96 \cdot SE \qquad SE = \frac{s}{\sqrt{n}},$$

[1, 2]

if a relative precision $E$ is required (i.e., the half-width of the interval is at most $E\bar{x}$), then

$$1.96\frac{s}{\sqrt{n}} \leq E\bar{x}.$$

Using $CV = \frac{s}{\bar{x}}$, this gives

$$n \geq \left(\frac{1.96 \times CV}{E}\right)^2.$$

Thus, the required number of trials increases with the square of the coefficient of variation: experiments with larger relative spread require more repetitions to achieve the same precision.

Usually, one of the first laboratory experiments is the vernier caliper measurement, where students are often asked to take six readings. One may ask: why six, and not three, five, or ten? The above statistical result provides the justification.

Suppose six measurements of a metal block thickness give a mean of 5.218 mm and a standard deviation of 0.327 mm, corresponding to a coefficient of variation of 6.27%. For a required relative precision of $E = 5\%$, the required number of trials is

$$N = \left(\frac{1.96 \times 0.0627}{0.05}\right)^2 = (2.456)^2 \approx 6.$$

Thus, for this level of variability, six vernier readings are statistically justified.

## 4 Is this universally constant?

The value $N \approx 6$ is not universal. The required number of trials depends on the variability of the system being studied.

To illustrate this, consider a percolation simulation (for example, water flowing through a porous medium). In percolation, flow occurs only if a continuous path spans the system. Because the placement of blocking particles is random, even a single blockage can completely stop the flow. This makes the outcome highly variable from trial to trial.

To study this variability quantitatively, we use a computational model of percolation implemented in NetLogo. NetLogo is an agent-based modelling platform in which individual elements of a system (such as open or blocked sites) are treated as independent agents that follow simple rules. We use the site-percolation model where each lattice site is randomly assigned as open or blocked with a given probability, and percolation occurs only if a connected cluster of open sites spans from one side of the system to the other. The global behavior of the system then emerges from these local interactions. By running the simulation many times under identical conditions, we can measure how much the outcome fluctuates from trial to trial.

A stochastic percolation simulation is performed in NetLogo, three particle sizes mimicking sandy clay is implemented in the percolation simulation. This produced a mean wet percentage of approximately 34% with a standard deviation of 10.29%, giving a coefficient of variation of 30.3%. This is much larger than in the vernier experiment (where $CV \approx 6.27\%$), indicating that the percolation outcomes fluctuate far more strongly from trial to trial. In other words, the system is intrinsically more variable.

Using the same relative precision $E = 5\%$, we obtain

$$N = \left(\frac{1.96 \times 0.303}{0.05}\right)^2 = (11.88)^2 \approx 141 \text{ trials.}$$

| Experiment | CV | $N$ ($E = 5\%$) |
|---|---|---|
| Wire (student) | 6.25% | **6** |
| NetLogo Wet% | 30.3% | **141** |

Table 2: Trials needed for standard 5% lab precision

Thus, for the same lab precision of $E = 0.05$ percolation requires many more trials because the system has much larger inherent variability. $N \propto (CV)^2$ explains the dramatic difference.

## 5 Conclusion

The required number of trials in an experiment is not a fixed or universal number. It depends on the intrinsic variability of the system being studied. Experiments with small relative fluctuations such as vernier calliper measurements require only a few repetitions to achieve a given precision. In contrast, highly stochastic systems such as percolation exhibit large trial-to-trial variability and therefore require many more repetitions to estimate the mean with the same confidence and precision.

In general, for a desired relative precision $E$ at 95% confidence, the required number of trials is given by

$$N = \left(\frac{1.96\,CV}{E}\right)^2.$$

This relation shows that the number of trials scales with the square of the coefficient of variation and increases rapidly for highly variable systems.

## References

[1] P.R. Bevington and D.K. Robinson, *Data Reduction and Error Analysis for the Physical Sciences*, 3rd ed., McGraw-Hill, 2003.

[2] J.R. Taylor, *An Introduction to Error Analysis: The Study of Uncertainties in Physical Measurements*, 2nd ed., University Science Books, 1997.

[3] C.M. Prestwood and K.V. Hartley, *Statistical Methods in Experimental Physics*, North-Holland, 1970.

[4] U. Wilensky, "NetLogo Percolation model," *Center for Connected Learning, Northwestern University, 1997*.