

On the Use of Random Sampling to Estimate the Value of π

Daniel Cellucci

*Department of Physics and Astronomy,
University of Georgia, Athens, Georgia 30602*

(Dated: January 25, 2012)

I. INTRODUCTION AND HISTORY OF THE PROBLEM

The exact value for the relationship between the diameter of a given circle and its circumference in flat space was an open problem to the ancient mathematicians. In particular, a problem called 'Squaring the Circle' was of special interest to geometers. They sought to answer whether, given a circle with a radius r , it would be possible to construct a square with an equal area using a finite number of steps and only a compass and straightedge. In 1882, the Lindemann-Weierstrauss theorem proved without any doubt that this venture was impossible by showing that π is a transcendental number, and is therefore only capable of being only estimated with a finite series.

An interesting fact: One needs only 39 digits of π to make a circle the size of the observable universe accurate to the size of a hydrogen atom.

II. BASIC PHYSICS

This lab seeks to "square" the circle, but instead of a compass and a straightedge, the tools being used will be a computer and "random" chance.

The methodology for such an estimation is straightforward. First, say the experimenter draws a circle at the bottom of a square bucket such that the center of the circle is coincident with the center of the bucket, and the circumference of the circle touches the sides of the bucket at four points. The experimenter then drops a series of balls into the bucket randomly. After dropping each ball, the experimenter then records whether the ball fell inside or outside the circumference of the circle (in the rare chance that a ball lands precisely on the circle, it is recorded as outside the circle). A running total of the number of balls to land inside circle is recorded, and the balls are removed each time to prevent balls from hitting one another. Once a satisfactorially large number of balls have been dropped, the experimenter then calculates the ratio γ , which is the number of balls that landed inside the circle over the total number of balls that were thrown in.

This ratio γ is related to the relative areas of the square bottom of the bucket and the circle inscribed inside. If the side of the bucket is a then the area of the bucket's square bottom is a^2 and the area of the circle is $\frac{\pi a^2}{4}$. The

ratio of these two areas is therefore equal to:

$$\frac{\pi}{4} = \gamma \quad (1)$$

III. DESCRIPTION OF THE SIMULATION

Though it might be satisfying to perform such an experiment in exactly the way described above, the computer can perform calculations such as these many orders of magnitude faster than the average human. However, the computer cannot generate genuinely random strings of numbers, and so it must instead rely on pseudo-random algorithms that can generate sequences of digits that are somewhat periodic. These periodicities are oftentimes so incredibly long that within a thousand iterations of a single program they will not begin to repeat and are thus, for all intents and purposes, random.

The first and most important part of this simulation, therefore, was the acquisition of a RNG that could handle a great many calls for random numbers. The experimenters utilized an RNG code common to the Center for Simulational Physics, called R1279.cc. The 1279 refers to the fact that, given a sufficiently large seed value, the RNG takes 10^{1279} calls before it begins to repeat itself. As this wasn't likely to be a very fruitful parameter to rigorously check, the experimenters instead settled to ascertain if a small sample of called values displayed the necessary randomness when graphed over the examined domain.

Figure 1 illustrates this effort. It is a plot of a small sample of the computational process employed by the computer, a set of random (x,y) coordinates graphed over a square area. All points inside the traced circle are counted toward the tally that will be used to estimate the value of π . As one can see from this sample, the points are effectively random in their distribution, with no noticable trends or patterns appearing.

Since the RNG was shown to be effectively random in its choice of parameters, the next step was to insure that this randomness continued throughout the generation of all of the values. As a result, the program called the RNG 2,000,000 times, then generated another value with which it used as a new seed. By changing the base generation value, this insured that the RNG continued to produce 'fresh' values, and avoid correlations that might otherwise ruin the accuracy of the method.

IV. MEASUREMENTS AND DISCUSSION

In all, 100 samples of 1,000,000 coordinates each were collected by the program. These values each produced an approximation of π that was randomly distributed around the true value. Figure 2 plots these values, as well as the mean and standard error.

These data produce a mean of 3.141591 and a standard error of 0.00017184. Thus, for these 100 points the actual value of π is well within the statistical error of the experiment. This corresponds to a 5.25×10^{-5} percent difference from the expected value of π , 3.14159265.

V. SUMMARY AND CONCLUSION

This experiment utilized chance and a simple algorithm to put a modern spin on an ancient problem- cal-

culating π using nothing but basic geometrical intuition. In all, over 200,000,100 random values were generated, and over 100,000,000 coordinates calculated to perform this "simple" calculation.

The results were worth it; an estimated value of π that is within 5.25×10^{-5} of the actual value. However, while this value is unquestionably accurate, there is a question of its precision. The value of 0.00017184 is two orders of magnitude larger than the absolute error between the value of π and its estimator. The only option to decrease this error is to make the sample set larger- doing so will decrease the size of the standard error by a factor of \sqrt{N} .

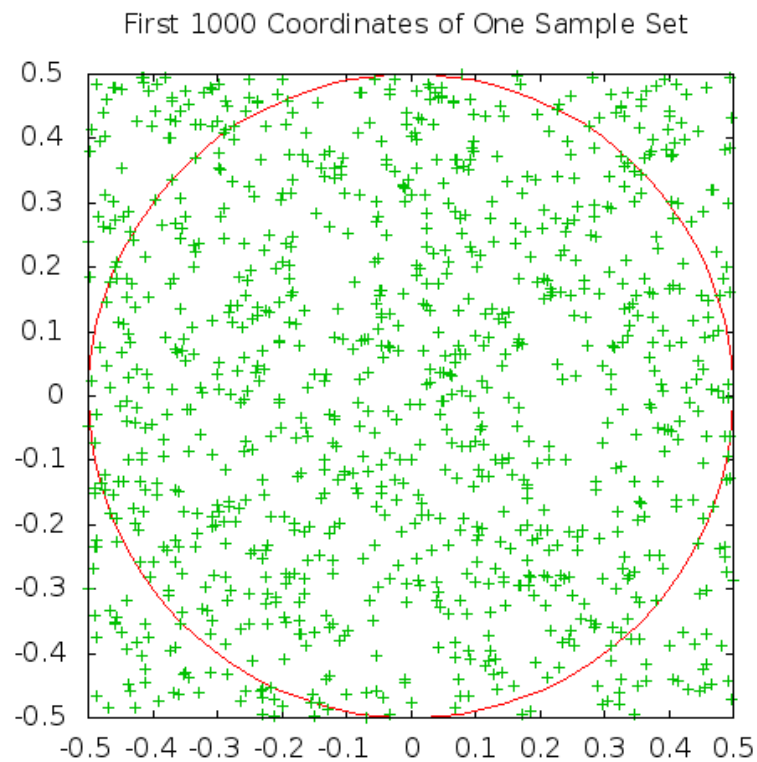


FIG. 1: Above is the graph of the first 1000 values of the first sample set. The green points are randomly generated coordinates, and the red is the boundary curve used to decide which points were inside the circle and which were outside.

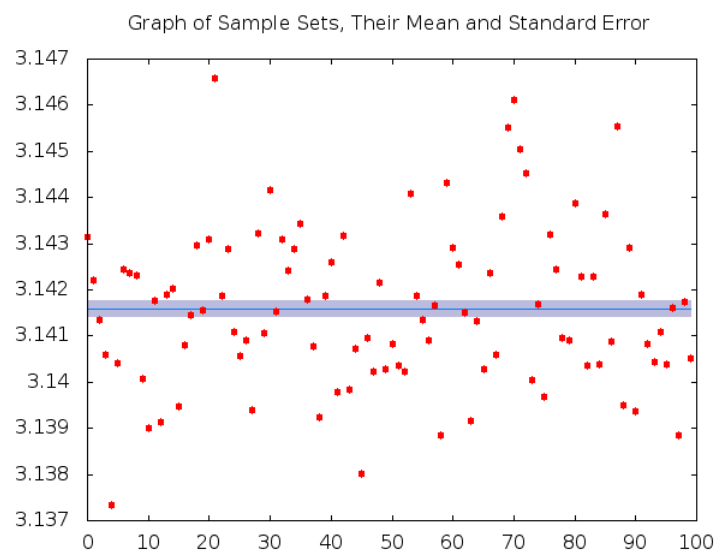


FIG. 2: The 100 values of the sample sets along with the average and the The value of π and the mean value estimator are so close in value that they are not discernible in this graph's range