Buffon's needles and the value of π

Awang Student (22BXXXX)

2022-10-26

Abstract

Consider a straight needle of length a, thrown 'randomly' onto an infinite plane covered with parallel lines spaced by distance d. A probabilistic argument yields that the probability of the needle intersecting with one of the lines involves the famous mathematical constant π. Through a series of "needle-tossing" experiments, the resulting data is analysed in order to estimate the value of π.

1   Introduction

In mathematics, the *Buffon's needle problem* is a question first posed in the 18th century by Georges-Louis Leclerc, Comte de Buffon:

> Suppose we have a floor made of parallel strips of wood, each the same width, and we drop a needle onto the floor. What is the probability that the needle will lie across a line between two strips?

Buffon's needle was the earliest problem in geometric probability to be solved [3]. The solution for the sought probability p, in the case where the needle length a is not greater than the width d of the strips, is

$$p = \frac{2a}{\pi d}. \tag{1}$$

This can be used to design a Monte Carlo method for approximating the number π, although that was not the original motivation for de Buffon's question.
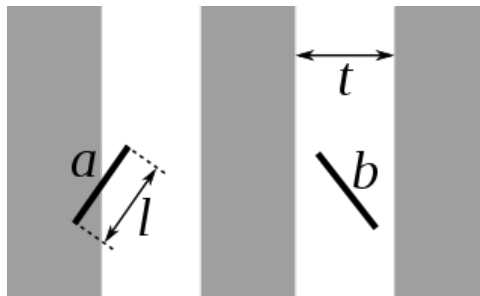

Figure 1: The a needle crosses the line, whereas the b needle does not. Source: Wikipedia

2   Derivation

The problem in more mathematical terms is: Given a needle of length a dropped on a plane ruled with parallel lines d units apart, what is the probability that the needle will lie across a line upon landing? Various sources exist for deriving the desired probability. See for example [2] and [1].

Consider the "short needle" case ($a < d$). For a randomly thrown needle, let θ be the angle that the needle makes between it and the horizontal lines, where $0 \le \theta \le \pi/2$. Note that the cases when the needle makes a 'negative slope', i.e. $\theta > \pi/2$ can be ignored because these can be argued to be identical to cases where $\tilde{\theta} = (\pi - \theta) \in [0, \pi/2]$ vis-à-vis the intersection of the needle with the horizontal lines.
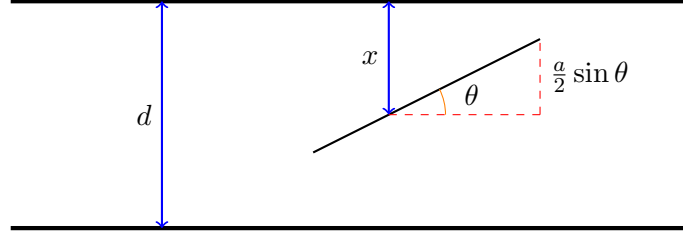

Figure 2: A randomly thrown short needle on a horizontal plane.

Let x be the shortest distance between the centre of the needle and a horizontal line. Thus $0 \le x \le d/2$. Using standard trigonometry, the *vertical half-height of the needle*, as measured between the centre of the needle to the "top point" of the needle closest to one of the horizontal lines, is given by $(a/2) \sin \theta$. **We realise that the needle crosses any horizontal line if $x < (a/2) \sin \theta$.**

As the needles are thrown randomly on the plane, values of θ and of x are *uniformly random*. Also, the values of θ and x do not influence one another, and thus are independent of each other. Recall that a random variable Y distributed as $\text{Unif}(\alpha, \beta)$ has pdf given by $f(y) = 1/(\beta - \alpha)$, where y takes values in $[\alpha, \beta]$. Therefore, we may write $\theta \sim \text{Unif}(0, \pi/2)$ with pdf $f(\theta) = 2/\pi$, and $x \sim \text{Unif}(0, d/2)$ with pdf $f(x) = 2/d$. Additionally, the bivariate random variable $\mathbf{z} = (\theta, x)$ is also uniformly distributed, with pdf

$$f(\mathbf{z}) = f(\theta, x) = f(\theta)f(x) = \frac{4}{\pi d},$$

$0 \le \theta \le \pi/2$, $0 \le x \le d/2$. The probability of crossing (p) is given by integrating the pdf over the set $\{(x, \theta) \mid x < (a/2) \sin \theta\}$, i.e.

$$
\begin{aligned}
p &= \iint\limits_{\{x < (a/2)\sin\theta\}} \frac{4}{\pi d} dx d\theta \\
&= \int_{\theta=0}^{\pi/2} \int_{x=0}^{(a/2)\sin\theta} \frac{4}{\pi d} dx d\theta \\
&= \frac{4}{\pi d} \int_{\theta=0}^{\pi/2} (a/2) \sin\theta d\theta \\
&= \frac{2a}{\pi d}
\end{aligned}
$$

Alternatively, we can obtain the probability by viewing the problem geometrically. Each point represents exactly one realisation of the position of the randomly landed needle. The points which represent needles crossing the line are shaded, and these are exactly the points in the set $\{(x, \theta) | x < (a/2) \sin \theta\}$. The area of the shaded region is obtained by integrating the function $f(y) = (a/2) \sin y$ from $y = 0$ to $y = \pi/2$, giving us $a/2$. Now the total area of the state space is the rectangular area $\pi/2 \times d/2 = \pi d/4$. Therefore, the probability required is

$$p = \frac{\text{Area of shaded region}}{\text{Total area}} = \frac{2a}{\pi d}. \tag{2}$$

3 Estimating π

Suppose that the value of p is estimated by a suitable value $\hat{p}$. Then the value π can then be estimated too by rearranging the equation in (1):

$$\hat{\pi} = \frac{2a}{\hat{p}d}. \tag{3}$$

Then the remaining question is how to obtain an estimate $\hat{p}$. From a frequentist perspective, one would be tempted to actually throw a bunch of needles and record how many times the needle actually crosses the line. Let B be the total number of times needles were thrown. The required estimator would then be

$$\hat{p} = \frac{\# \text{ times cross the line}}{B}.$$

Indeed, such an estimator would also be guaranteed to converge to the actual true value p as $B \to \infty$ by the weak law of large numbers.

It was said that Buffon actually performed this experiment not with needles, but with baguettes thrown from a tower onto the city square below. Perhaps slightly less excitingly, we will perform the simulations in R code, and this will be explained in the next section.

4 Simulations

There are two determinants of the way the needles are dropped: the location of the midpoint of the needle along the y-axis (call this `y`); and the angle that the needle makes with the horizontal axis (call this `angle`). Assuming independent and uniform throws of the needle, `y` can be simulated using `runif()` with `min = -d/2` and `max = d/2`, where `d` is defined to be the distance between two lines. This makes it such that the 0 line is in the middle, and the closest distance x between the needle midpoint and a horizontal line is obtained by `d/2 - abs(y)`. Note that the `x` position along the horizontal axis is not too important here; but for nice results we simulate x uniformly from `runif(min = 0, max = 2 * d)`.

The `angle` may be also be uniformly distributed, and the minimum and maximum value for this variable is 0 and π respectively. Depending on the value of `angle`, the corresponding pair of end points for the randomly thrown needle can be obtained using basic trigonometry, deduced from the midpoint position x and y. That is, the upper and lower end-point in the y-axis direction is `y + (a/2) * sin(angle)` and `y - (a/2) * sin(angle)`, irrespective of the value of `angle`. The left and right end-point in the x axis direction meanwhile is `x - (a/2) * sin(angle)` and `x + (a/2) * sin(angle)` for when `angle` is less than $\pi/2$; if `angle` is greater than $\pi/2$ then these values are reversed.

To run the simulation, values of `a` and `d` have to be chosen such that `a < d`. Here, we chose `a <- 3` and `d <- 4`. We need to also choose some value `B`, noting that the bigger this number is, the better the results. We chose `B <- 2000`. Figure 3 below plots the first 250 iterations of the needle simulation. The needles have been colour coded according to whether or not the horizontal lines were crossed.

Of course, the main interest is in the estimate $\hat{\pi}$. The value of $\hat{\pi}$ is tracked as the simulation progresses (a running estimate) and plotted against the simulation number. The results are show in Figure 4. The value $\hat{\pi}$ is seen to get closer to the actual value of π (horizontal dotted line) as the simulation increases. After 2000 simulations, the value of π was estimated to be 3.168, giving an absolute error value of 0.026. The estimate for first 500 needle tosses seem to be quite poor, suggesting that if this experiment was done in actuality, a really large sample size would be required.
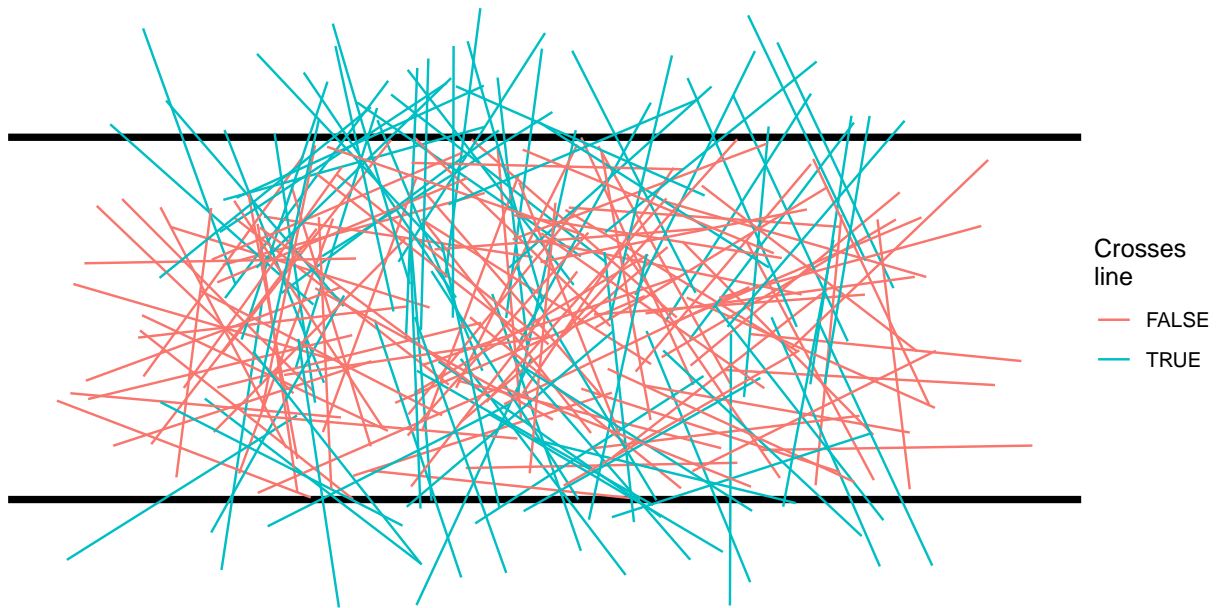
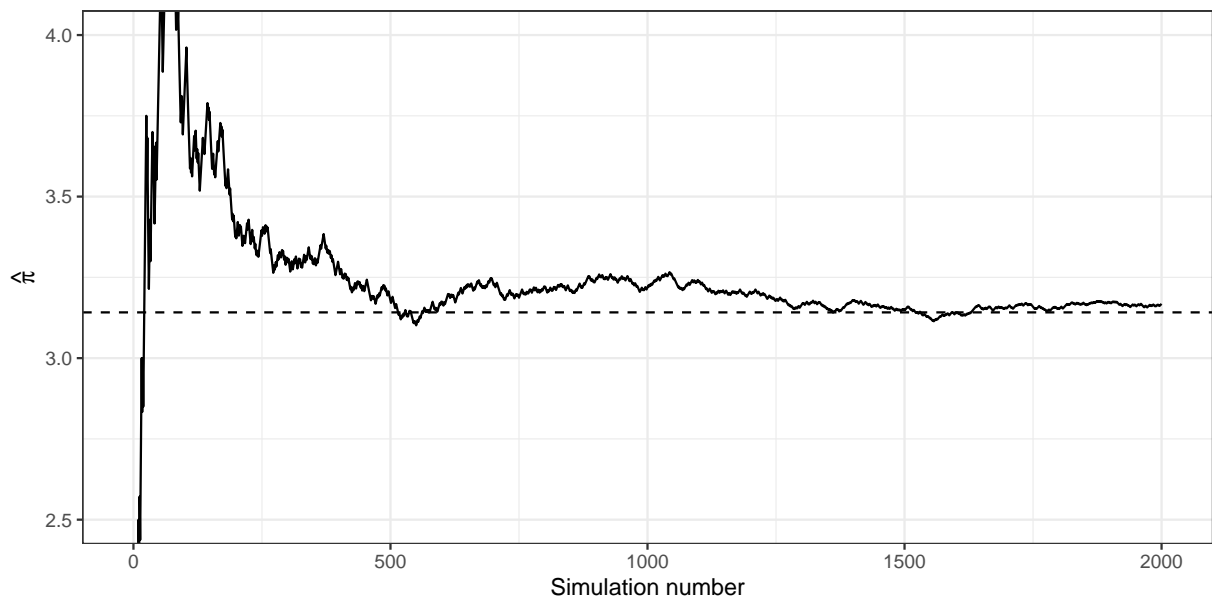Figure 3: The first 250 random throws of the needles.

Figure 4: The running estimate for the value of π.

5 Conclusions

A computer-based simulation of the Buffon's needle problem was presented. For a simulation size of 2000, the value of π was estimated to be 3.168, giving an absolute error value of 0.026. In the course of conducting this mini project, further questions were pondered:

1. What values of a and d are optimal in terms of minimising the error for predicting π?

2. What needs to be changed, if at all, for the "long needle" scenario, i.e. for values $a > d$?

Both seemingly interesting questions which are hoped to be tackled another time.

A   Code for simulation

```
set.seed(261022)
B <- 2000  # Number of simulations
d <- 4     # Distance between horizontal lines
a <- 3     # Length of needle

# Create the tibble
plot_df <-
  tibble(
    throw = 1:B,
    x = runif(B, 0, 2 * d),
    y = runif(B, 0 - d/2, d/2),
    angle = runif(B, 0, pi),
    x1 = case_when(
      angle > pi/2 ~ x + a/2 * cos(pi - angle),
      TRUE ~ x - a/2 * cos(angle)
    ),
    y1 = y - a/2 * sin(angle),
    x2 = case_when(
      angle > pi/2 ~ x - a/2 * cos(pi - angle),
      TRUE ~ x + a/2 * cos(angle)
    ),
    y2 = y + a/2 * sin(angle),
    s = d/2 - abs(y),  # shortest distance between midpoint and hor. line
    theta = case_when(  # theta value in [0,pi/2]
      angle > pi/2 ~ pi - angle,
      TRUE ~ angle
    ),
    cross = s < a/2 * sin(theta),  # does the needle cross?
    pi_pred = 2 * a / (cummean(cross) * d)  # running estimate for pi
  )

# Estimate for pi
pihat <- 2 * a / (mean(plot_df$cross) * d)
```

B   Code for figures

B.1   Needle simulation figure

```
plot_df %>%
```

```
slice_head(n = 200) %>%
ggplot() +
geom_hline(yintercept = c(-d/2, d/2), size = 1.5) +
geom_segment(aes(x1, y1, xend = x2, yend = y2, group = throw, col = cross)) +
coord_equal() +
theme_void() +
labs(col = "Crosses\nline")
```

B.2   Running estimate figure

```
ggplot(plot_df, aes(throw, pi_pred)) +
  geom_line() +
  geom_hline(yintercept = pi, linetype = "dashed") +
  theme_bw() +
  coord_cartesian(ylim = c(2.5, 4)) +
  labs(y = expression(hat(pi)), x = "Simulation number")
```

References

[1]   RL Duncan. "A variation of the Buffon needle problem". In: *Mathematics Magazine* 40.1 (1967), pp. 36–38.

[2]   Nathan Mantel. "An extension of the Buffon needle problem". In: *The Annals of Mathematical Statistics* 24.4 (1953), pp. 674–677.

[3]   Eric W Weisstein. "Buffon's Needle Problem". In: (2003). URL: https://mathworld.wolfram.com/.