Buffon’s Needle Problem

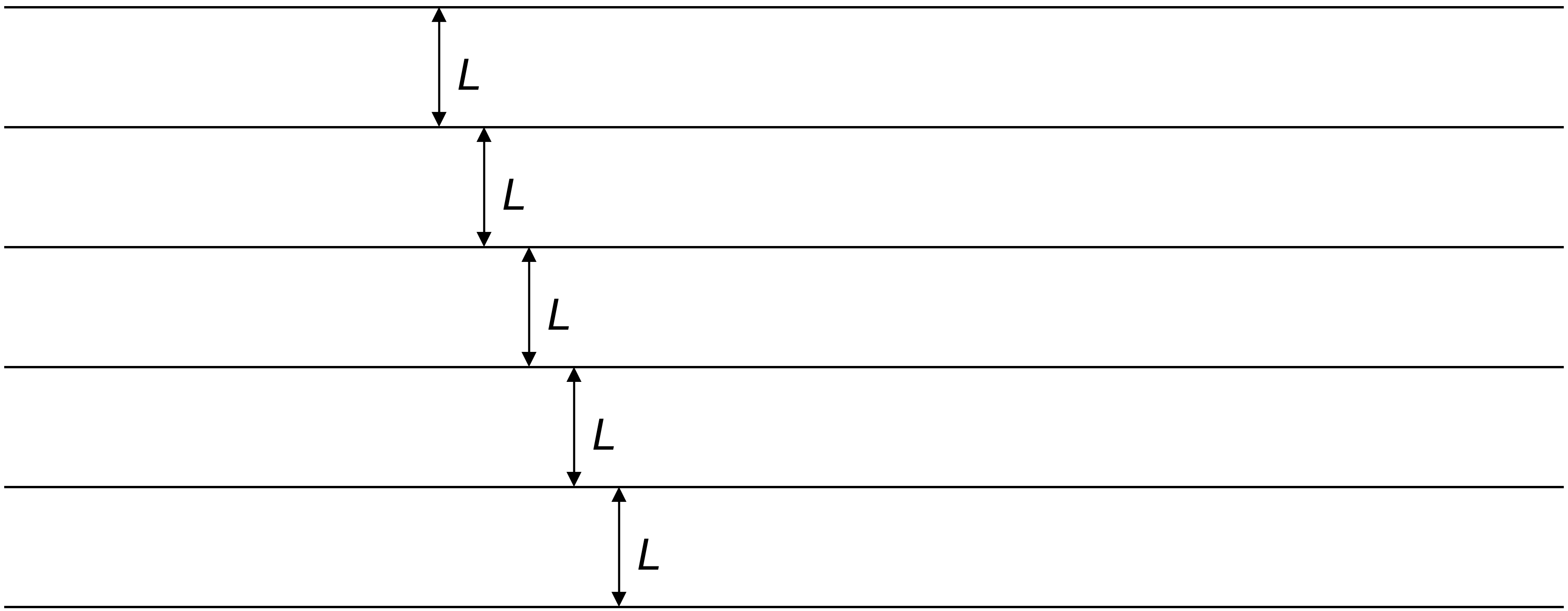
Clay Ford
October 15, 2010

Backstory

- Georges-Louis Leclerc, Comte de Buffon
- Known primarily for his contributions to Natural history (35 volumes) in 18th century
- Posed his needle problem in a paper published in 1777

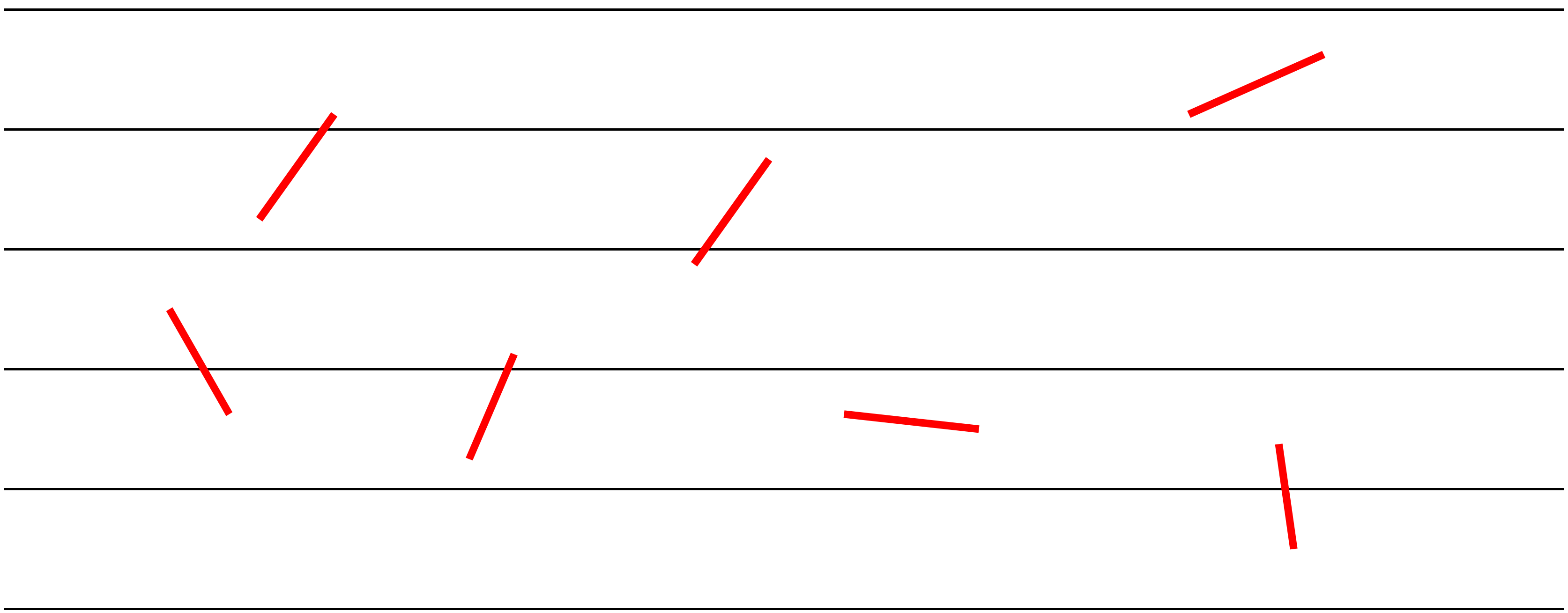
Problem Setup

- A floor with parallel lines, all an equal distance apart (call it *L*):



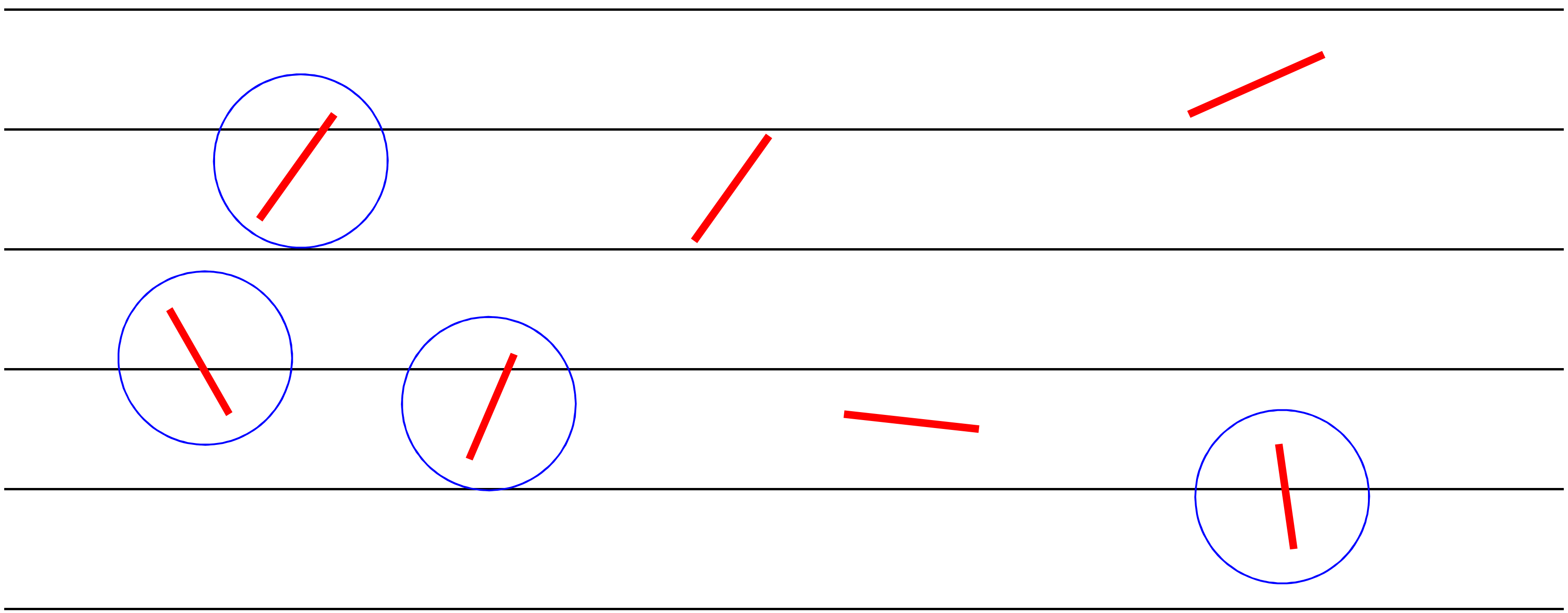
Problem Setup

- Imagine we drop some needles with length $a = L$ on the floor:



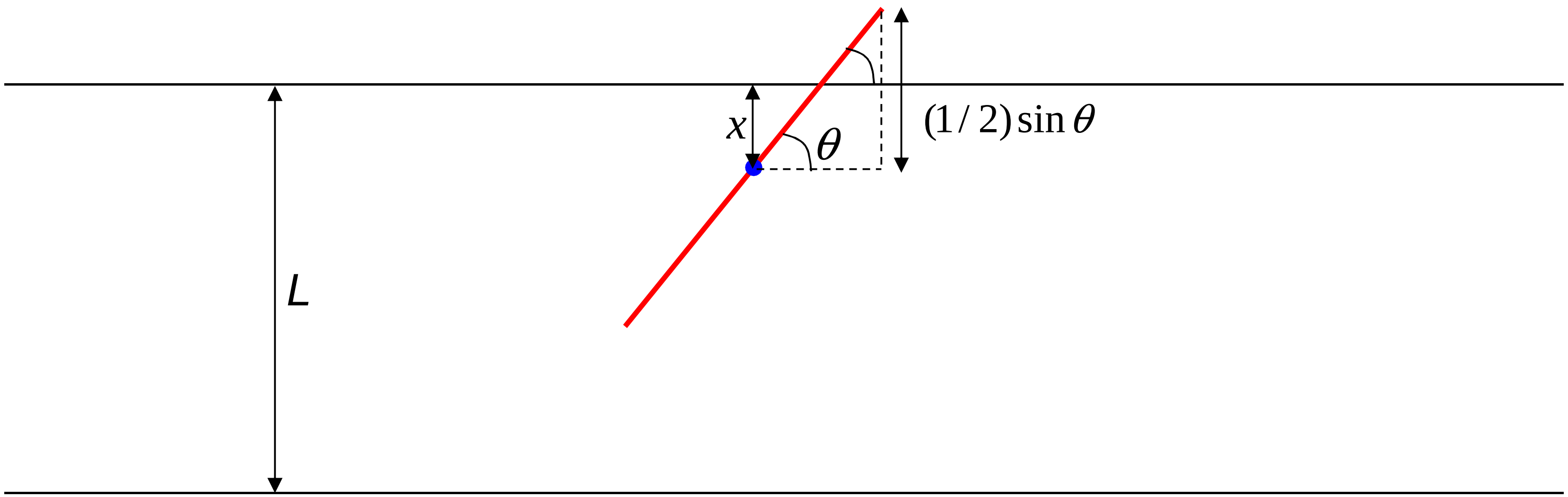
Problem Setup

- What's the probability a needle intersects a line?



When does a needle intersect a line?

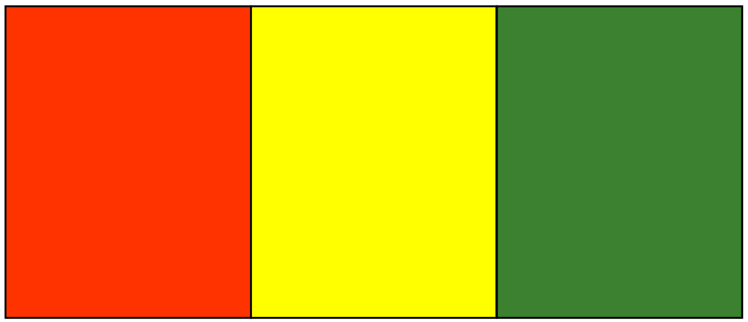
- When $x \le (1/2)\sin\theta$, where x is distance from midpoint of needle to closest line.



For simplicity, we let $L = a = 1$

What's the probability?

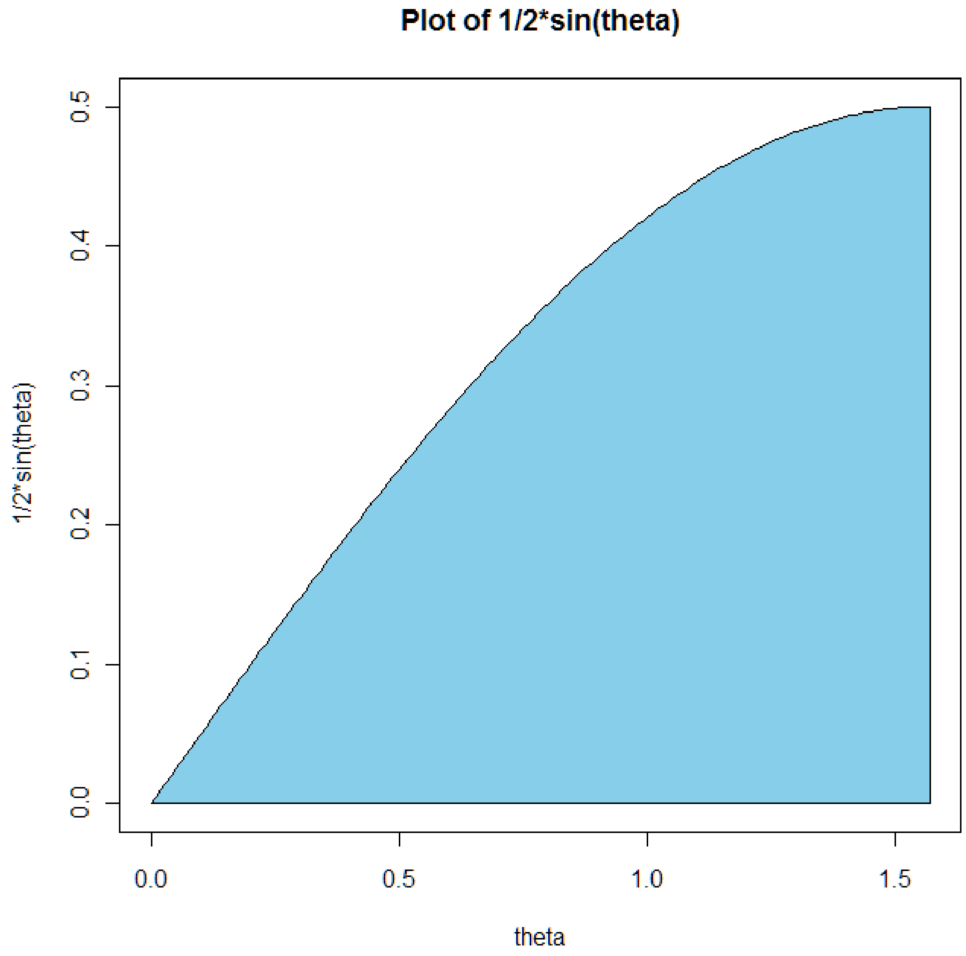
- We seek $P(x \le (a/2)\sin\theta)$
- Frame the problem in terms of geometry
- For example:



$$P(\text{landing on green}) = \frac{\text{green area}}{\text{total area}}$$

Define a region

- For simplicity, set $L = a = 1$
- Plot $(1/2)\sin\theta$ as a function of θ
- $0 \le \theta \le \pi/2$
- The region in blue represents when $x \le (1/2)\sin\theta$

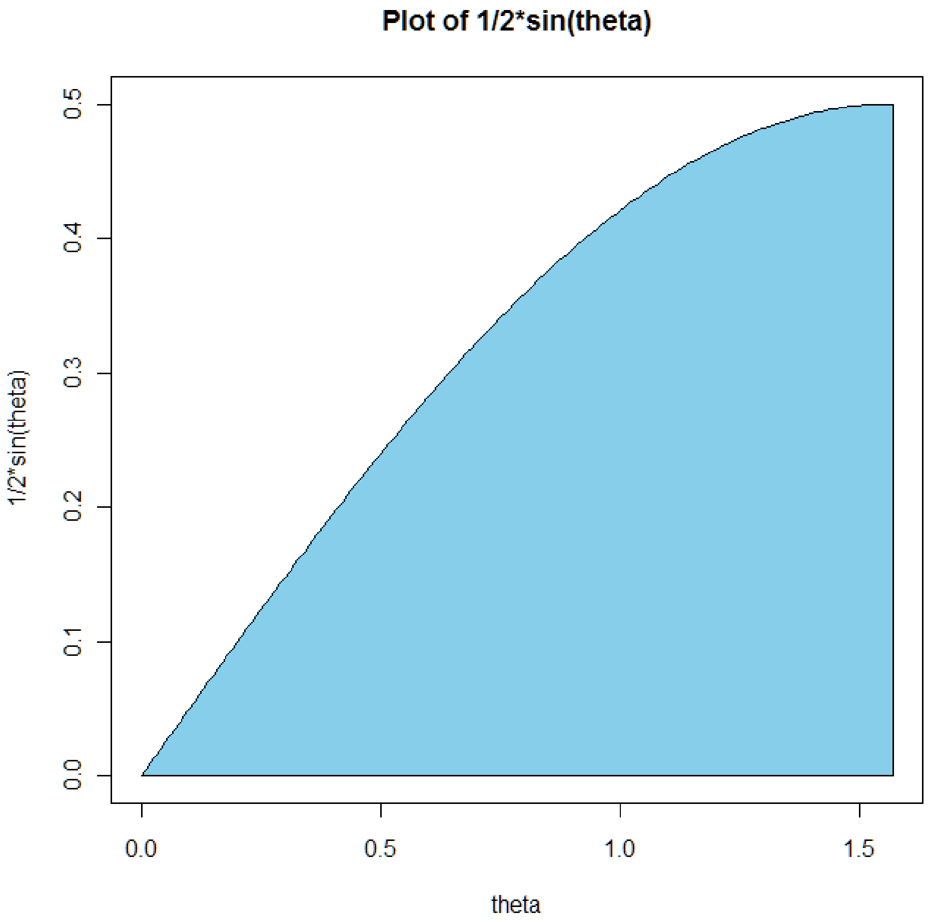


Area of blue region

$$\int_0^{\frac{\pi}{2}} \frac{1}{2}\sin\theta \, d\theta$$

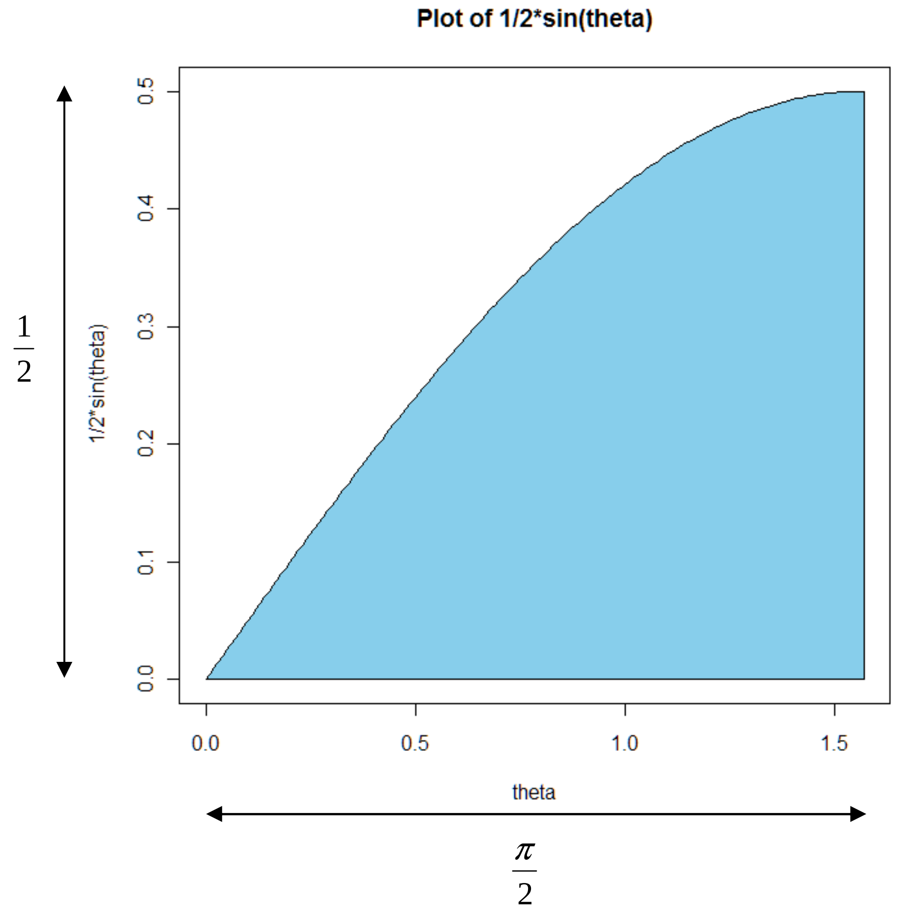
$$= \frac{1}{2}\left[-\cos\frac{\pi}{2} - (-\cos 0)\right]$$

$$= \frac{1}{2}$$



Area of total region

$$A = \frac{1}{2}\frac{\pi}{2} = \frac{\pi}{4}$$



Now find probability

$$P(x \le (1/2)\sin\theta)$$

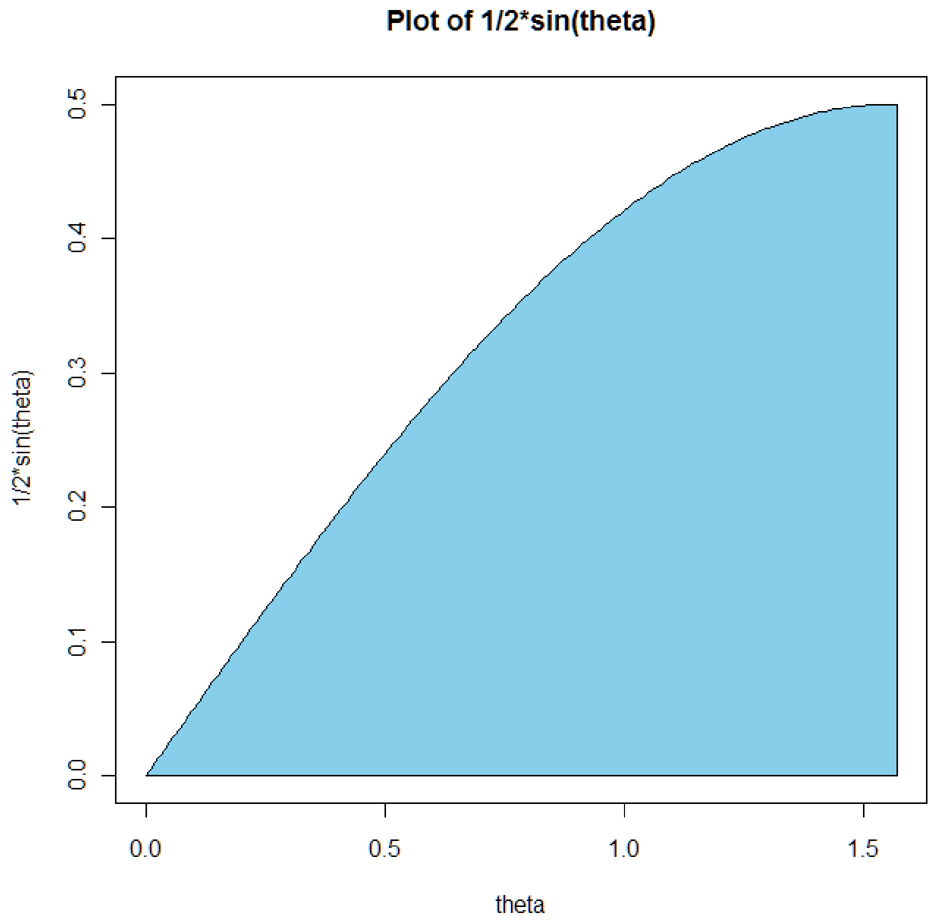
$$= \frac{\frac{1}{2}}{\frac{\pi}{4}} = \frac{1}{2}\frac{4}{\pi}$$

$$= \frac{2}{\pi} \approx 0.6366197$$

In general, we can show

$$P(x \le (a/2)\sin\theta) = \frac{2a}{L\pi}$$

for $a \le L$



Let's Drop Some Needles

- Again, assume $a = 1$ and $L = 1$
- P(line intersection) = $2/\pi$
- The process of dropping *n* needles and observing intersections has the Binomial setting
  - Fixed number of trials
  - Independence
  - Same probability on each throw of a needle
  - Two outcomes: cross a line/do not cross a line

Expected number of crossings

- Under Binomial setting, expected number of successes (i.e., line intersections) is

$$E(X) = np$$

- Example: say we toss 1000 needles
  - Expected number of line crossings is 1000 × 2/π
  - That's about 637

Estimating probability

- Further, we could toss 1000 needles and estimate *p* if we didn't know it
  - Say we toss 1000 needles and observe 650 hits
  - $\hat{p} = \frac{x}{n} = \frac{650}{1000} = 0.65$

The implication

- But let's pretend we knew *p* but did not know the value for π
- Notice we can solve for π:

$$\hat{p} \approx p = \frac{2}{\pi} \approx \frac{650}{1000}$$

$$\Rightarrow \pi \approx \frac{2n}{\text{number of hits}} = \frac{(2)1000}{650} \approx 3.08$$

- The frequency of needles crossing lines can be used to estimate π!

Simulation

- Using R or SAS, we can create a program that simulates needle tosses and estimates π
  - randomly generate an acute angle
  - randomly generate the distance of midpoint of needle to nearest line
  - If $x \leq (1/2)\sin\theta$ then record a line crossing

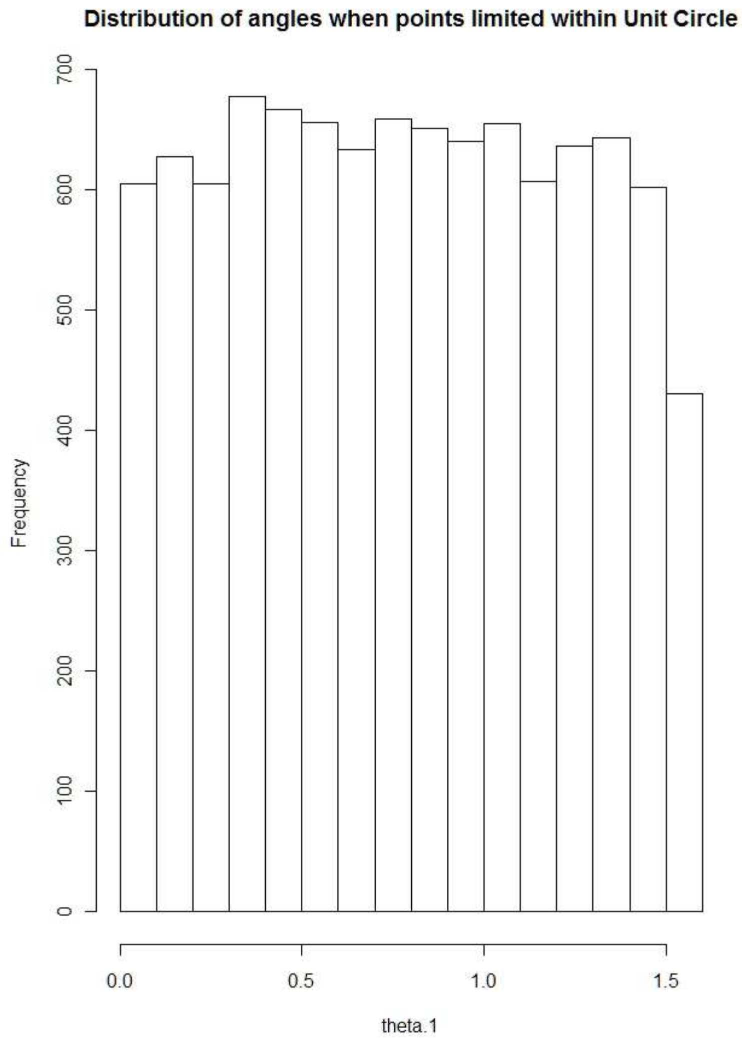
Simulation Code - R

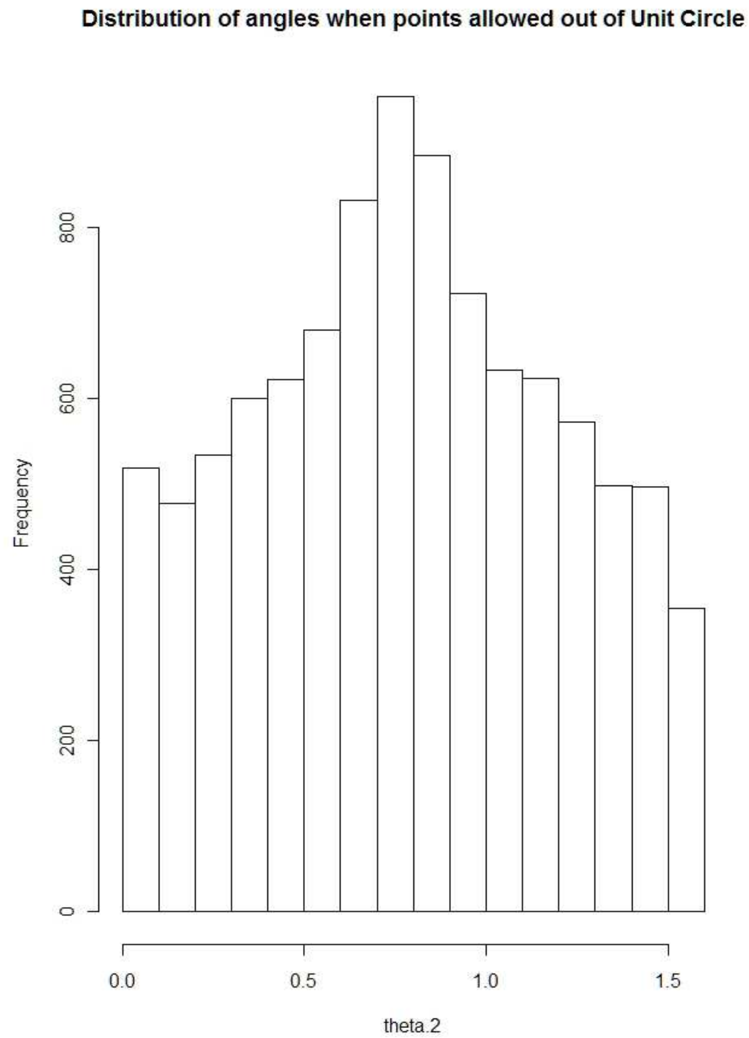
```
a <- 1 # length of needle
L <- 1 # distance between lines
n <- 100000 # number of dropped needles
hit <- 0
for(i in 1:n) {
        x <- runif(1,0,1)
        y <- runif(1,0,1)
         while(x^2 + y^2 > 1) { # no points outside of unit circle
                x <- runif(1,0,1)
                y <- runif(1,0,1)
        }
        theta <- atan(y/x) # the random angle
        d <- runif(1,0,(L/2)) # distance of needle midpoint to
                                      nearest line
        if(d <= (a/2)*sin(theta)) {
                hit <- hit + 1
        }
}
pi.est <- (n*2*a)/(hit*L)
```

Simulation Code - SAS

```
data _null_;
file log;
hit = 0;
a = 1; /*length of needle*/
L = 1; /*distance between lines*/
n = 1000000;
do i = 1 to n;
    x = ranuni(0);
    y = ranuni(0);
        do while (x**2 + y**2 > 1);
            x = ranuni(0);
            y = ranuni(0);
        end;
    theta = atan(y/x);/*random angle*/
    d = (a/2)*ranuni(0);/*distance from middle of needle to nearest line*/
if d <= (L/2)*sin(theta) then hit = hit + 1;
end;
pi_est = (n*2*a)/(hit*L);
put 'estimate of pi: ' pi_est;
run;
```

Why keep points in Unit Circle?





Simulation Results ($a = L$)

- For $a = 1$, $L = 1$, $n = 10{,}000$
  - 3.148119
- For $a = 1$, $L = 1$, $n = 100{,}000$
  - 3.140605
- For $a = 1$, $L = 1$, $n = 1{,}000{,}000$
  - 3.141680

First 6 digits of pi:
3.141592

Simulation Results ($a < L$)

- For $a = 1$, $L = 1.25$, $n = 10{,}000$
  - 3.175863
- For $a = 1$, $L = 1.25$, $n = 100{,}000$
  - 3.134735
- For $a = 1$, $L = 1.25$, $n = 1{,}000{,}000$
  - 3.139847

First 6 digits of pi:
3.141592