# Lecture 16: Buffons Needle and CLT

Junaid Hasan

Math 381

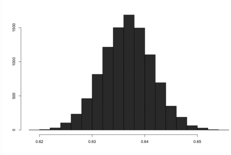
# Buffons Needle and Central Limit Theorem

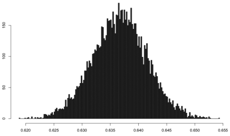
- An important theorem in probability is the Central Limit Theorem. It says (roughly) that if we sample many values from an unknown distribution and take their mean, and do this multiple times, the means will be approximately **normally distributed**.
- A normal distribution (also known as **Gaussian**) has the shape of $y = e^{-x^2}$ with scaling.
- Features: It is symmetric about its peak, it has exactly two inflection points, and is asymptotic to the x-axis.

# contd..

- Suppose we throw the needle in Buffon's needle N $=1000$ times.
- We can view this as sampling from \${0,1 } bernoulli trials with 1 meaning the needle crosses the line and 0 meaning it does not.
- Thus throwing 1000 times and calculating empirical probability can be seen as averaging 1000 random samples from $\{0,1\}$ with the probilities given.
- If we do this many times, and make a histogram of the results we will notice a normal shape.

# Experiment

- Running M multiple runs of N = 1000 needle throws and recording the estimated probability from each run yeilds something like
- 

- 

# CLT lessons

- The CLT says what these histograms suggest that the distribution of these values is very nearly normal.
- Furthermore we can use statistics now to get confidence intervals.

# CLT Application

- Let us say $M = 1000$ and the values are $x_1, \ldots, x_{1000}$.
- The mean is
$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i = 0.6364$$
- The standard deviation was
$$s = 0.0158$$

.

# Contd..

- For a true normal 99.73% of the values are within 3 standard deviations of the mean.
- Hence we can be very confident that the true probability $p$ lies between $\bar{x} - 3s$ and $\bar{x} + 3s$. In other words we can be 99.73% confident that

$$0.5892 < p < 0.6837.$$

- If our confience is 95% then the value must be within 2 standard distrubutions

$$0.6049 < p < 0.6680.$$

# Contd..

- We saw that with 99.73% confidence we can claim that

$$0.5892 < p < 0.6837.$$

- What if we want sharper bound?
- Either we must reduce the confidence (not desirable)
- Or we must increase $M$ or $N$.
- Which one should we increase?

# Answer

- We must make each run with more steps, increase $N$.
- Notice this is different from the number of runs M.
- Number of runs M does not affect standard deviation significiantly, since each sample mean still is obtained the same steps.

# Getting confidence interval

- Say we run our simulation 10 times and get 10 estimates.
- If each estimate is a mean of the same distribution, then by CLT we know that these 10 estimates are normally distributed.
- Normal means that the data must be **symmetric** about mean, therefore the probability that the estimates are on one side of the true probability is

$$\frac{2}{2^{10}} \approx 0.00195 \approx 0.2$$

# Contd..

- Therefore the probability that all of the estimates are greater, or all of the estimates are less than the true probability is 0.2.
- Say we run the Buffon Needle $M = 10$ times with $N = 10^6$.
- I got [0.6375, 0.6365, 0.6363, 0.6361, 0.6366, 0.6360, 0.6370, 0.6367, 0.6362, 0.6362]
- Therefore the true probability lies in the interval $[0.6361, 0.6375]$ with 99.8% calculation.
- Note that this method does not require standard deviation computation.