Niraj Mahajan | Raaghav Raaj

Page No.:
Date:

CS215: Assignment 4

Question 1: Estimating π.

* Bivariate Random Variable,

$$X := (X_1, X_2)$$

X_1, X_2 independent and $X_1, X_2 \sim U(-1, 1)$

(a) ~~The~~ For X to lie inside a circle of radius 1, centered at origin,

$$X_1^2 + X_2^2 \leq 1$$

Hence, the area of the preferred region on plane,

$$A_1 = \pi \times (1)^2$$

$$A_1 = \pi$$

Now,

the area of the region on which the X varies is,

$$A_2 = 2 \times 2$$

$$A_2 = 4$$

~~At~~ The probability that X lies inside the unit circle, is given as,

$$P = \frac{|S_1|}{|S_2|},$$

where S_1 is the set of ~~elements~~ values of X which lie in a circle.

and,

S_2 is the set of all values of X.

Clearly,

the ratio of the cardinality of S_1 and S_2 is proportional to the ratio of the areas.

Hence, $P = \frac{A_1}{A_2}$

$\Rightarrow P = \frac{\pi}{4}$

Also, the probability of the above event can be found in the following manner.

$$P = \int_{-1}^{1}\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{2}\cdot\frac{1}{2}\, dy \cdot dx. \quad , \text{ where } x = X_1 \text{ and } y = X_2.$$

$$\Rightarrow \boxed{P = \int_0^1 \sqrt{1-x^2}\cdot dx = \frac{\pi}{4}}$$

(b) The above question brings out the fact that,

$$\pi = 4 \times P\left((X_1, X_2) \text{ lying in circle}(0,1)\right)$$

For the above estimation, we draw N sample points $X := (X_1, X_2)$, where

$X_1 \sim \text{uniform}(-1, 1)$

$X_2 \sim \text{Uniform}(-1, 1)$

Let us define a random Bernoulli variable, $M(X(x_1, x_2))$ such that,

$$M := \begin{cases} 1 & \text{, when } X \text{ lies inside the circle} \\ 0 & \text{, otherwise} \end{cases}$$

Thus, the probability,
P(X lies inside the circle) equals.

$$\frac{\sum_{i=1}^{N} M(X_i)}{N}$$

Hence, the estimated value of π can be given as,

$$\pi_{estimate} = 4 \times \frac{\sum_i M(x_i)}{N}$$

(c) Usage of code.

The following are the instructions for the usage of the code:

- Load the code present in 'submission/code/q1/q1.m
- Run the code in 'q1.m'
- The output resulted is:

$N = 10$	$\pi_{est} = 2.8000$
$N = 100$	$\pi_{est} = 2.9200$
$N = 10^3$	$\pi_{est} = 3.1600$
$N = 10^4$	$\pi_{est} = 3.1388$
$N = 10^5$	$\pi_{est} = 3.1453$
$N = 10^6$	$\pi_{est} = 3.1414$

$N = 10^7$ $\pi_{est} = 3.1416$

$N = 10^8$ $\pi_{est} = 3.1412$

For the values of N larger than 10^9, we can use the following method.

- Run the loop $\frac{N}{10^6}$ times, for the code of $N' = 10^6$, where N' is the sample set count, and take the average over N.
  - → This allows us to use only 10^6 amount of memory, but 1000 times, for $N = 10^9$, which is managable.

Our code handles this case and outputs the value of π_{est} as 3.1416.

i.e. $N = 10^9$ $\pi_{est} = 3.1416$

(d) As per the question, we wish to estimate the value of π in the range $[\pi - 0.01, \pi + 0.01]$.

Obviously, for such small errors, the value of N must be large.

Thus, using the central limit theorem for large N, the random variable π_{est} can be approximated as a Gaussian distribution.

So we have,

- $$E(\pi_{est}) = 4 \times E\left[\frac{\sum_i M(x_i)}{N}\right]$$

$$= \frac{4}{N} \sum_i E[M(x_i)] \quad (x_i : \text{iid variables})$$

$$= \frac{4}{N} \times N \times \frac{\pi}{4}$$

$$= \pi$$

- $$\mathrm{Var}(\pi_{est}) = 16 \times \mathrm{Var}\left[\frac{\sum_i M(x_i)}{N}\right]$$

$$= \frac{16}{N^2} \times \mathrm{Var}\left[\sum_i M(x_i)\right]$$

$$= \frac{16}{N^2} \times N \times \cancel{4}\, \mathrm{Var}[M(x_i)] \quad (\text{iid variables})$$

$$= \frac{16}{N} \times \frac{\pi}{4}\left(\cancel{\pi}\; 1 - \frac{\pi}{4}\right) = \frac{\pi(4-\pi)}{N} = \sigma^2$$

Hence, $\pi_{est} \sim G\left(\pi, \frac{\pi(4-\pi)}{N}\right)$.

We look up the tables for normal distribution and find that 95% of data lies in

$$(\mu - 1.96\sigma, \mu + 1.96\sigma).$$

In the question,

$$1.96\sigma \leq 0.01$$

$$\Rightarrow \sigma \leq \frac{1}{196}$$

$$\Rightarrow (196)^2 \times \pi \times (4 - \pi) \leq N.$$

$$\therefore \boxed{N \geq 103599}.$$