Day 3 → Oct 5, 2024

Understanding & Overview of Probability & Statistics in Machine Learning

What is probability?

→ A mathematical tool used to understand, handle, study and analyze the uncertainty and randomness

→ handle uncertainties and study real-world randomness phenomena

Figure Gaussian | Normal Distribution

Equation of Quantum System showing uncertainty due to inherent stochasticity

$$ih \frac{\partial}{\partial} \psi(x,t) = \left[-h^2 \frac{\partial^2}{} + V(x,t) \right.$$

$$\frac{\partial t}{\partial x^2}$$

$$J\Psi(x,t)$$

where probability $= |\Psi|^2$

→ Incomplete Observation,

→ Information or modeling

Deep Learning Training:.

$$E_{n(x,u)}[\Gamma(L_{(m)}..)]$$

$\mathbb{E}_{p(x,y)}[\mathcal{L}(f_\theta(x), y)]$

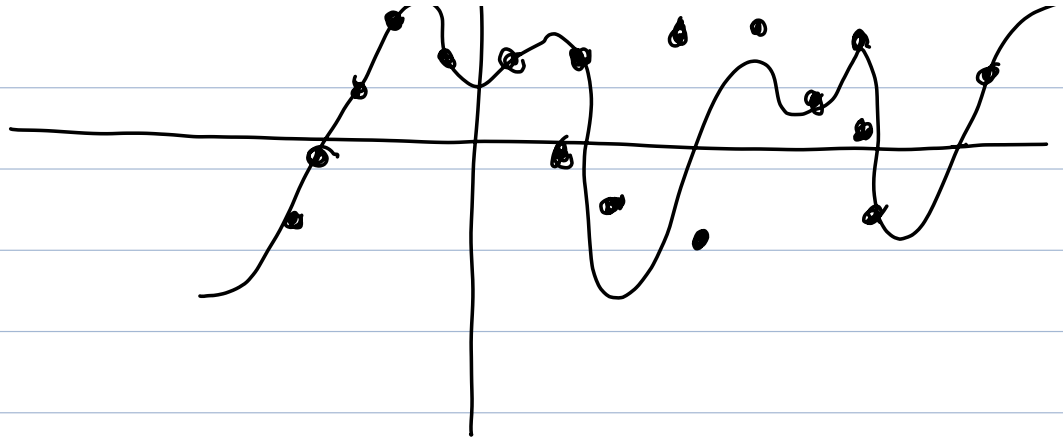
Expectation wrt $p(x,y)$

$\mathcal{L}$ → Loss Function

f_θ → prediction and y is Label

$f\theta$ is ANN parameters

prediction with uncertainty



Incorporating uncertainty in the real world interaction

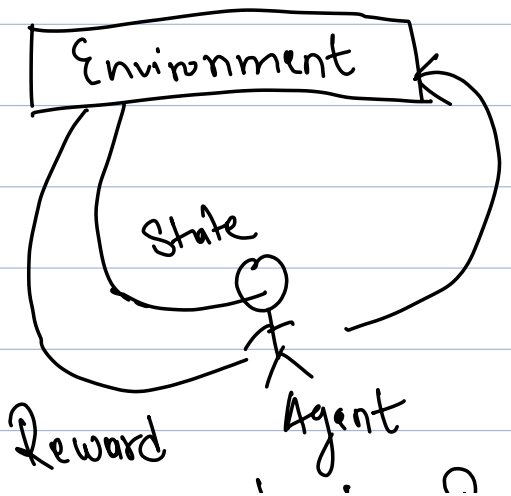


Fig. showing Reinforcement Learning

→ Environment is not static, filled with uncertainty
→ Agents does action, triggers events in environment

Basic Probability

* Random Variable - coin toss, rolling of dice,

X = 1 head and x = 0 tail

where we donot Know the probability of 'x' occuring or the value of x.

Examples: position of electron, Price of Stock,

getting trial or 3 run in a poker game.

* Framework to set probability

→ with value 0 to 1 we can set 'p' probability

◦ Sample Space

X can either be 0 or 1

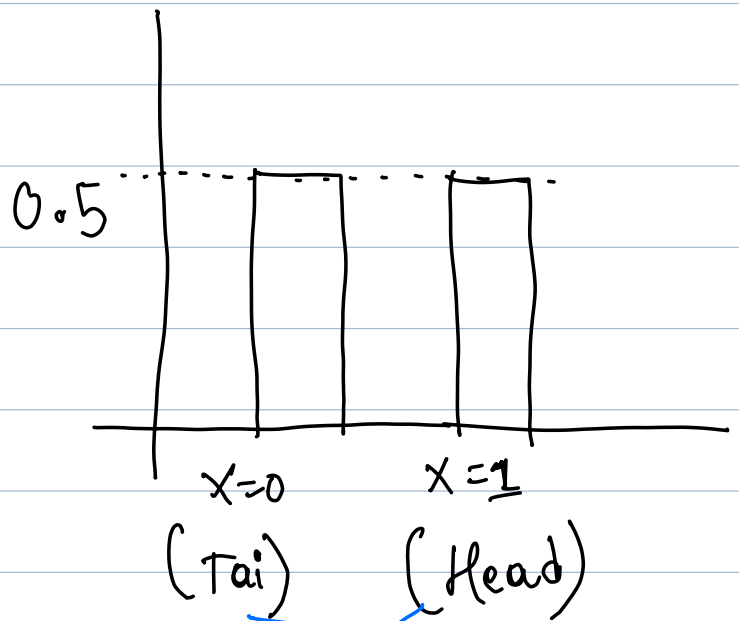
$p(x=0) = 0.5$

$p(x=1) = 0.5$

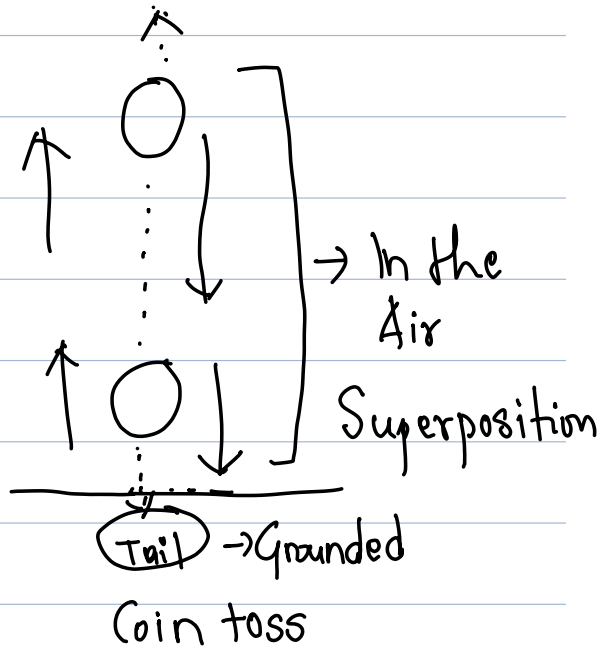
$\Sigma = p(x=0) + p(x=1)$

$\Rightarrow 1$ (Sum of p is 1)

Mapping the discrete or Continuous event in a range of 0 and 1 is probability.



(state Space)



Statistics (Inferring From Data)

* Suppose 5 times Coin toss and the results are

1, 0, 1, 0, 1

$p(x=1) \Rightarrow 0.8$

$p(x=0) \Rightarrow 0.2$

So, Statistics is the inferring from Data with parameters (variables) and underlying process explains data, evaluation of data and the process like tossing coss output

Distributions

→ describe the probability

→ description of probability of all events in sample space

→ discrete and continuous depends on sample space

data type

* Discrete (10, 30, 20)
* Continuous (1.53, 9.89, 23.56%)

Distributions plays crucial role to describe and define different probability for different events.

Bernoulli Distribution

Eg:- Coin toss have only two States 0 and 1 and it has parameter ϕ

So,

$P(x = 1) = \phi$

$p(x=0) = 1 - \phi$

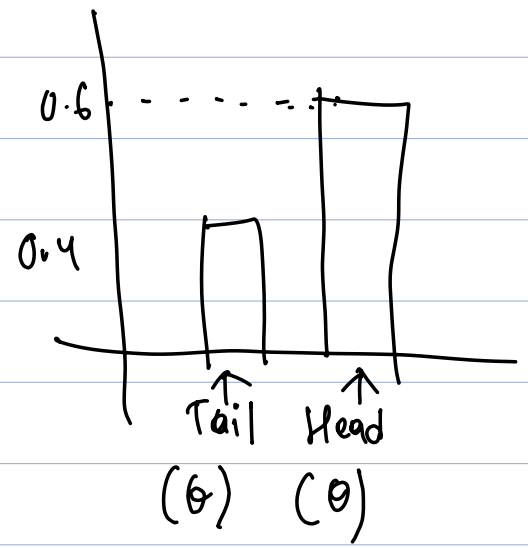
$p(x=x) = \phi^{x}(1-\phi)^{1-x}$ (PMF)

Expectation $\rightarrow E_x[x] = \phi$, $Var_x(x) = \phi(1-\phi)$ (variance)

Parameter (ϕ)

using Bernoulli Distribution,

we can use PMF (and need not to explicitly mention, tail this or Head (p) this)



PMF (Probability Mass Function) describes the probability

of each possible event

$$P(X=x) = \phi^{x}(1-\phi)^{1-x}$$

put I want 'p' 1 then $x=1$

$$P(X=1) = \phi^{1}(1-\phi)^{1-1}$$

$$\Rightarrow \phi$$

put for $x=0$

$$P(x=0) = \phi^{0}(1-\phi)^{1-0}$$

$$\Rightarrow 1-\phi$$

this is a way to describe the 'p' in a more

mathematical and concise way.

PMF (Probability Mass function)

A function P is PMF if

$$\forall x \in \mathbb{X}: 0 \leq p(x) \leq 1$$

∴ the integral exists

$$\sum_{x} p(x) = 1 \rightarrow p \text{ must be sum to '1'}$$

The domain $\mathbb{X}$ contains all possible states x can take.

$\forall$ → there exists

x → possible states taken

$0 < p(x) < 1$ between or equal to 0 and

less than or equal to 1

0 to 1 (you can take any Range but 0 and 1 is easy and the standard way)

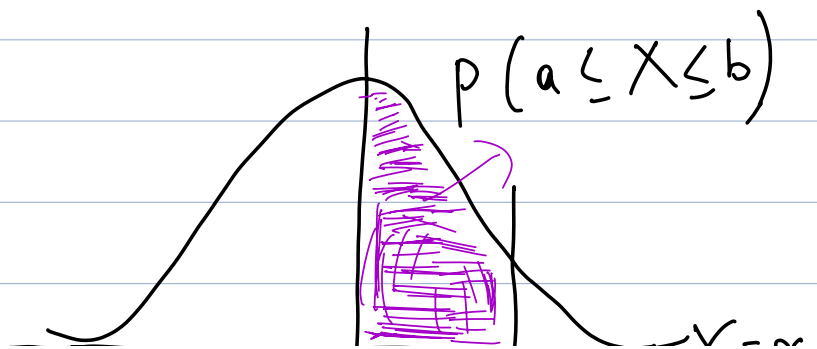
Continuous Distributions

→ X takes continuous value, X is random variables

→ PDF instead of PMF

So, the probability lies in between a and b

$p(a \leq X \leq b)$



if we need to find the 'p'.

Anamoly / Unusual
→ throwing dart

→ Random Variable takes continuous value in a range

in between $-\infty$ to $+\infty$ and finding or getting $p = 1$ that is very absrud / never hit the $p = 1$.

P ⇒ event between a and b
instead of p(a) i.e. 0

∫ Area → $p(a \leq x \leq b)$ events