# Seaborn

Vizualization Distribution
library that uses Matplotlib to plot graphs.
Used to vizualize random distribution.

Displots
Stands for distribution plots.
Takes array as input.
plots a curve corresponding to the distribution points.

Note Matplotlib is used to do so.

```
import matplotlib.pyplot as plt
import seaborn as sns

sns.displot([0, 1, 2, 3, 4])
plt.show()
```

if we do not want histogram the we can use parameter 'Kind = "kde"'

```
sns.displot([0, 1, 2, 3, 4], kind = 'kde')
```

i Normal Distribution
Most important distribution
Also called Gaussian distribution.
Fits probability of many events eg → IQ score, heartbeat.

we use random.normal() method.

Three parameters.

- ~~me~~ loc → (Mean) where peak of bell exists.
- scale → (standard deviation) how flat the graph should
- size → shape of returned array

Note loc and scale are not mendatory.

```
import numpy as np
from numpy import random as np.

x = np.random.normal(loc=1, scale=2, size=(2,3))
print(x)
```

⇒ creates a 2x3 distribution.

Note curve of normal distribution also known as bell curve.

ii) Binomial distribution
It is a Discrete distribution.
describe outcome of binary scenarios ie. toss of coin etc.

Three parameter

n - number of trials
p - probability of each trial
size - shape of array.

Discrete Distribution
seperate set of events eg toss of coin

While height of people is continous.

```
from numpy import random
x = random.binomial(n=10, p=0.5, size=10)
print(x)
```

=> [4 7 3 2 8 6 3 4 7 3]

Here, For each value (10 values) the success is counted out of 10 trials.

For first value '4', it denotes their were 4 succesful trials out of 10.

Note: if there are enough data points it will be quite similar to normal distribution with certain loc and scale.

iii Poisson Distribution
Discrete distribution.
estimates how many times an event can happen in specified time.

Two parameters

lam → Known no. of occurence
size → shape of array.

```
x = random.poisson(lam=2, size=10)
print(x)
```

PAGE NO.:

Normal vs Poisson

- Normal is continous while Poisson is discrete.
- For large enough poisson value it will become similar to normal

Binomial vs Poisson

Binomial can have only two possible outcome while poisson can have number of outcomes.

For very large n and near-zero p, it is near identical to poisson such that n*p is nearly equal to lam.

```
binomial = random.binomial(n=1000, p=0.01, size=1000)
poison = random.poisson(lam=10, size=1000)
```

iv) Uniform distribution

describe probability where every event has equal chance of occuring. Eg generation of random numbers.

Three parameters

low - lower bound, 0.0 default
high - higher bound, 1.0 default.
size - shape of array.

i) continous uniform distribution

eg → generation of real number.

```
x = random.uniform(low=2, high=5, size=5)
print(x)
```

Teacher's Signature........................

⇒ generate 5 floating numbers between 2 and 5.

2 Discrete uniform distribution
have given limited values. eg rolling a dice.
Here numbers 1-6 have equal probability in all.

x = random.randint(low=2, high=5, size=5)

V logistic Distribution
used to describe growth
used extensively in ML in logistic regression etc.

Three parameters:

loc → mean, Default 0
scale → standard deviation. default 1
size → shape of array.

logistic vs Normal
- Near identical but logistic has more area under tails.
- represents more probability of occurence of an event further away from mean.

Note:- higher value of scale, normal is identical apart from peak.

PAGE NO.:

## viii Multinomial distribution

- generalization of binomial distribution.
- describe probabilities of multinomial scenarios. eg blood type

Three parameters

n - no of trails
pvals - list of probability outcomes.
size - shape.

Note Multinomial samples will not produce a single value. They will produce one value for each pval.

x= random.multinomial(n=6, pvals[1/6]* 6])

## viii Exponential Distribution

used in describing time till next event eg failure/success

Two Parameter

scale - inverse of rate. default is 1.0.
size - shape.

Relation Between Poisson & Exponential

poisson deals with number of occurences of an event in a time period.
exponential distribution deals with time between these events.

PAGE NO.:

viii) Chi square Distribution
Used as a basis to verify the hypothesis

Two parameters

df - degree of freedom
size - shape

x = random.chisquare(df=2, size=(2,3))

sns.distplot(x)
plt.show()

ix) Rayleigh Distribution
Used for signal processing. It has two parameters:-

scale - decide how flat distribution will be default 1.0
size - shape of array.

Note: At unit std and 2 deg of freedom rayleigh and chi square represent the same distribution

x) Pareto Distribution
Pareto's law ie 80-20 distribution (20% factor cause 80%)
Two parameter

a - shape parameter
size - shape

Teacher's Signature........