Discrete Distributions

- ↳ Bernoulli Distribution
- ↳ Binomial Distribution
- ↳ Uniform "
- ↳ Poisson "
- ↳ Geometric "

Binomial Distribution

A binomial random variable is the number of successes x in n repeated trials of a binomial experiment. The probability distribution of a binomial random variable is called a binomial distribution.

(i) A binomial experiment is a statistical experiment that has the following properties:

(i) The experiment consists ~~of~~ of n repeated ~~trails~~. trials.

- Each trial can result in just two possible outcomes — heads or tails.
- The trials are independent; that is, getting heads on one trial does not affect whether we get heads on other trials.

Notation

- The following notation is helpful, when we talk about binomial probability.
- x: The no. of successes that result from the binomial experiment.
- n: The number of trials in the binomial experiment
- P: The probability of success on an individual trial
- Q: The probability of failure on an individual trial. (This is equal to $1-P$).
- $n!$: The factorial of n (also known as n factorial)
- $b(x; n, P)$: Binomial probability — the probability that an n-trial binomial experiment results in exactly x successes, when the probability of success on an individual trial is P.
- ${}_nC_r$: The number of combinations of n things taken r at a time.
- Binomial Formula: Suppose a binomial experiment consists of n trials and results in x successes. If the probability of success on an individual trial is P, then the binomial probability is

$$b(x; n, P) = {}_nC_x * p^x * (1-P)^{n-x}$$

$$= \frac{n!}{(n-x)!} * p^x * (1-P)^{n-x}$$

- Expected value = mean = np
- Variance = npq

① Flip a coin 10 times. Let X number of heads obtained.

② A worn machine tool produces 1% defective parts. Let x no. of defective parts in the next 25 parts produced.

③ Each sample of air has a 10% chance of containing a particular rare molecule. Let X the number of air samples that contain the rare molecule in the next 18 samples analyzed.

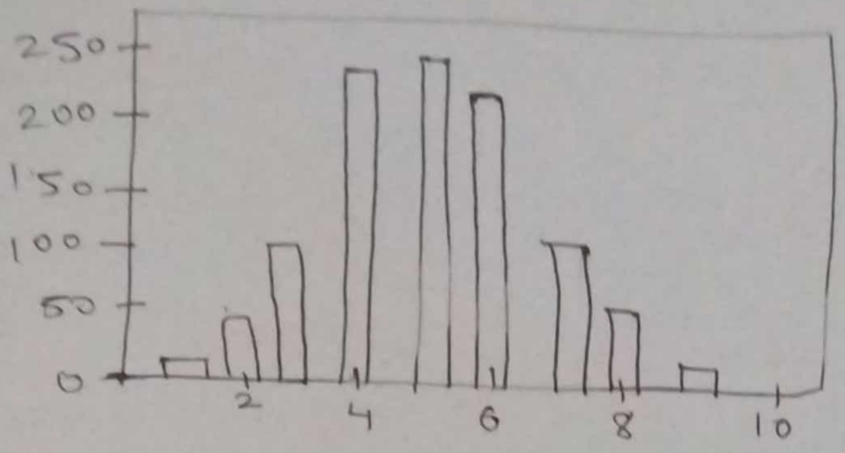
④ Of all bits transmitted through a digital transmission channel, 10% are received in error. Let x the number of ~~questions answered correctly.~~ bits in error in the next five bits transmitted.

⑤ A multiple choice test contains 10 questions, each with four choices, and you guess at each question. Let X the no. of questions answered correctly.

⑥ In the next 20 births at a hospital, let X the no. of femal births.

⑦ Of all patients suffering a particular illness, 35% experience improvement from a particular medication. In the next 100 patients administered the medication, let x the no. of patients who experience improvement.

```
from numpy import random
import matplotlib.pyplot as plt
import seaborn as sns
x = random.binomial(n=10, p=0.5, size=10)
print(x)
sns.distplot(random.binomial(n=10, p=0.5,
             size = 1000),
plt.show()
```



Bernoulli Distribution.

- Bernoulli distribution is easiest distribution among all distributions.
- It is similar to binomial distribution. The only difference is it takes only one trial while binomial distribution considers n trial.
- It has only two possible outcomes i.e success vs ~~family~~ failure.
- Let's consider random variable X with only one parameter p which represents probability of occurrence of event.

It's density function is given as:

$$P[X=1] = p$$
$$P[X=0] = 1-p$$

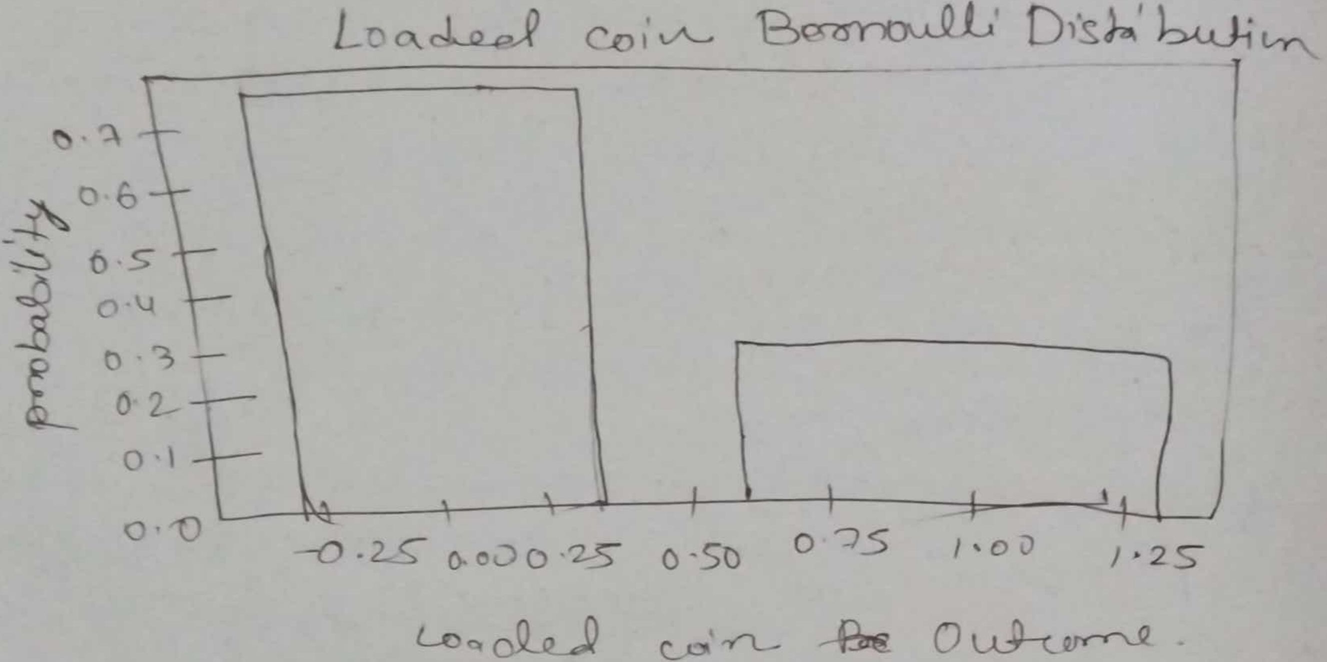
where,

$X=1$ indicates event has occurred.
$X=0$ " " didn't "

$$E[X] = p$$

$$Var[X] = p(1-p)$$

```
import numpy as np
import matplotlib.pyplot as plt
p = np.array([0.75, 0.25])
face = [0,1]
plt.bar (face, p)
plt.title ('Loaded coin Bernoulli Distribution')
plt.ylabel ('Probability')
plt.xlabel ('Loaded coin Outcome')
plt.show( )
```



Note

Bernoulli ~~theorem~~ Distribution & Binomial Distribution both are ~~same~~ gives same result but only one different is ~~Dis~~ Bernoulli Distribution check only one time and Binomial check many time.

Uniform Distribution.

- There are two kinds of uniform random variables: discrete and continuous ones.
- A discrete uniform distribution will take a (finite) set values S, and assign a probability of 1/n to each of them, where n, is the amount of elements in S.

- This way, if for instance my variable Y was uniform in $\{1, 2, 3\}$, then there'd be a 33%. chance each of those values came out.
- A very typical case of a discrete uniform random variable is found in dice, where your typical dice has the set of values $\{1, 2, 3, 4, 5, 6\}$.

- A random variable X has a discrete uniform distribution if each of the n values in it's range, say $x_1, x_2, \ldots, x_n$, has equal probability.

Then, $f(x_i) = 1 \quad n$

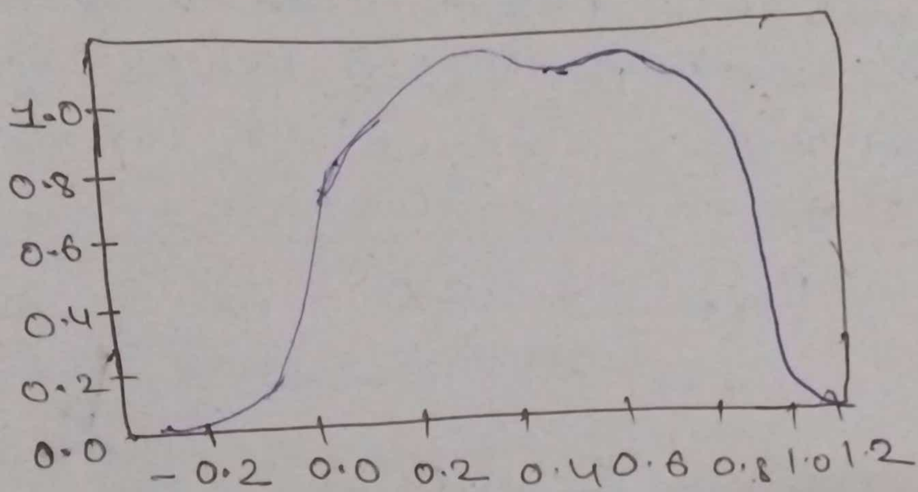
$$f(x) = \begin{cases} 1/n, & \text{if } x \in X \\ 0, & \text{Otherwise} \end{cases}$$

. 2. Continuous uniform distribution

. Not all uniform distributions are discrete; some are continuous. A continuous distribution (also referred to as rectangular distribution) is a statistical distribution with an infinite number of equally likely measurable values. Unlike discrete random variables, a continuous random variable can take any real value within a specified range.

- A continuous uniform distribution usually comes in a rectangular shape. A good example of a continuous uniform distribution is an idealized random number generator. With continuous uniform distribution, just like discrete uniform distribution, every variable has an equal chance of happening. However, there is an infinite number of points that can exist.

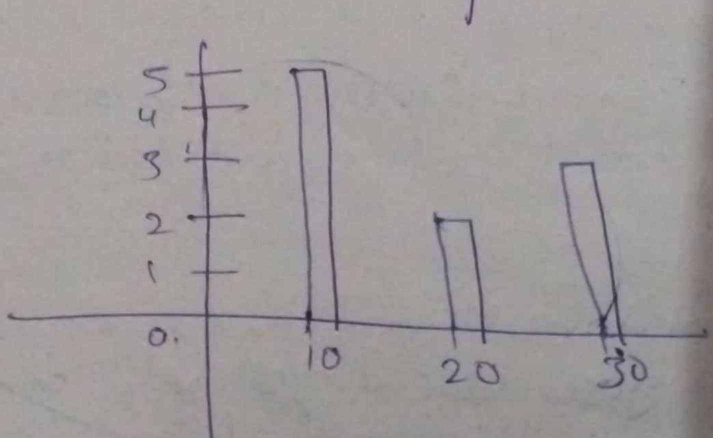
```
from numpy import random
import matplotlib.pyplot as plt
x=random.uniform(size=(2,3))
print(x)
sns.distplot(random.uniform(size=1000),hist=
                                       False)
plt.show()
```



Histogram

1 5 18 19 2 3 21 22 29 9

(0 to 10) → 5 | (11 to 20) → 2 | 21-30 → 3



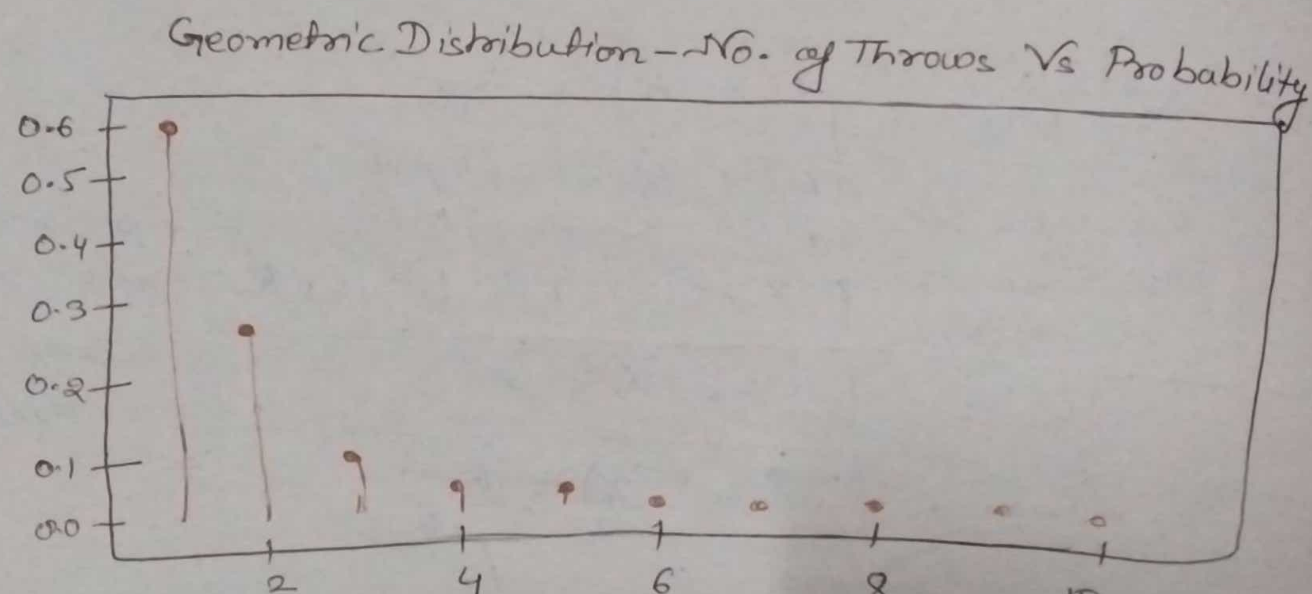
Geometric Distribution

A geometric distribution is defined as discrete probability distribution of a random variable 'x' which statisfies some of the conditions. The geometric distribution conditions are

- A phenomenon that has series of trials.
- Each trial has only two possible outcomes — either success or failure.
- The probability of success is the same for each trial.
- In probability and statistics, geometric distribution defines the probability that first success occurs after k number of trials. If p is the probability that success occurs on the k^th. ~~trial~~ Trial is given by the formula

Probability mass function	$P(X=k)=p(1-p)^{k-1}$	
Cumulative Distribution function	$p(X\leq k)=1-(1-p)^4$ $p(X\geq k)=(1-p)^{k-1}$ $p(X>k)=1-P(X\leq k)=(1-p)^k$	p = probability of success k = # of trials
Mean	$\mu=E(X)=\frac{1}{p}$	
Variance	$\sigma^2=V(X)=\frac{(1-p)}{p^2}$	

```
from scipy.stats import geom
import matplotlib.pyplot as plt
X=[1,2,3,4,5,6,7,8,9,10]
p= 0.6
geom_pd = geom.pmf(X,p)
fig, ax = plt.subplots(1,1, figsize=(8,6))
ax.plot(X,geom_pd, 'bo', ms=8, label='geom pmf')
plt.ylabel('Probability', fontsize='18')
plt.xlabel('X- No. of Throws', fontsize='18')
plt.title('Geometric Distribution - No. of Throws
        Vs Probability', fontsize='18')
ax.vlines(X,0, geom_pd, colors='b', lw=5,
                 alpha=0.5)
```



Poisson Distribution

- Poisson distribution is discrete probability distribution.
- Poisson distribution is a distribution of count ie number of times event has occured in given interval of time
- Poisson distribution can be used to predict probability of number of successful event that

may occur in specific interval of time.

⊙ Example, if a call center recieved 50 calls in 1 hour, then using Poisson distribution we can predict probability of getting 20 calls in next 30 minutes.

The random variable X that equals the number of counts in the interval is a Poisson random variable with parameter.

Probability Mass Function	$P(X) = \frac{e^{-\lambda}\lambda^x}{x!}$
Mean	$\mu = E(X) = \lambda$
Variance	$\sigma^2 = V(X) = \lambda$
Standard Deviation	$\sigma = \sqrt{\sigma^2} = \sqrt{\lambda}$

$0 \le x \le \infty$

λ is the average number of occurences in an interval.

```
from numpy import random
import matplotlib.pyplot as plt
import seaborn as sns
```

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

```
sns.distplot(random.poisson(lam=2, size=1000), kde=False)

plt.show()
```