#Statistics Advance Assignment

Q1. What is the Probability density function?

A probability density function (PDF)

is a mathematical function that describes the likelihood of a continuous random variable taking on a particular value within a given range. The PDF represents the relative likelihood of different outcomes occurring within the range of possible values for the random variable.

Q2. What are the types of Probability distribution?

Probability distributions are categorized into several types, each with its own characteristics and applications. Here are some common types of probability distributions:

1 Discrete Probability Distribution:

Represents the probabilities associated with discrete or countable outcomes. Examples include the Bernoulli distribution, Binomial distribution, Poisson distribution, and Geometric distribution.

2 Continuous Probability Distribution:

Describes probabilities associated with continuous random variables.

Examples include the Normal (Gaussian) distribution, Exponential distribution, Uniform distribution, and Chi-Square distribution.

3 Bernoulli Distribution:

Represents the probability distribution of a random variable that takes the value 1 with probability p and the value 0 with probability 1−p.

It is used for modeling binary outcomes, such as success/failure or yes/no experiments.

4 Binomial Distribution:

Describes the probability distribution of the number of successes in a fixed number of independent Bernoulli trials.

It is characterized by two parameters: the number of trials n and the probability of success p.

5 Poisson Distribution:

Models the number of events occurring in a fixed interval of time or space, given a constant average rate of occurrence.

It is often used in scenarios involving rare events, such as the number of phone calls received by a call center in a given hour.

6 Normal (Gaussian) Distribution:

One of the most widely used probability distributions, characterized by a bell-shaped curve. It is symmetric and defined by two parameters: the mean μ and the standard deviation σ. Many natural phenomena tend to follow a normal distribution, making it useful in various fields, including statistics, finance, and natural sciences.

7 Exponential Distribution:

Describes the time between successive events in a Poisson process, where events occur continuously and independently at a constant average rate. It is characterized by a single parameter, the rate parameter

λ, which represents the average rate of event occurrences.

8 Uniform Distribution:

Represents a constant probability over a continuous interval. All outcomes within the interval are equally likely. It is often used when there is no particular bias towards any value within the interval.

Q3. Write a Python function to calculate the probability density function of a normal

distribution with given mean and standard deviation at a given point.

In [1]:
```
import math

def normal_pdf(x, mean, std_dev):
    """
    Calculate the probability density function (PDF) of a normal distributio
    at a given point x with the specified mean and standard deviation.

    Arguments:
    x (float): The point at which to evaluate the PDF.
    mean (float): The mean of the normal distribution.
    std_dev (float): The standard deviation of the normal distribution.

    Returns:
    float: The value of the PDF at the given point x.
    """
    coefficient = 1 / (std_dev * math.sqrt(2 * math.pi))
    exponent = -0.5 * ((x - mean) / std_dev) ** 2
    pdf_value = coefficient * math.exp(exponent)
    return pdf_value

# Example usage:
mean = 0
std_dev = 1
x = 1
pdf_at_x = normal_pdf(x, mean, std_dev)
print("PDF at x =", x, ":", pdf_at_x)
```

```
PDF at x = 1 : 0.24197072451914337
```

Q4. What are the properties of Binomial distribution? Give two examples of events where binomial distribution can be applied.

The binomial distribution is a discrete probability distribution that describes the number of successes in a fixed number of independent Bernoulli trials, where each trial has only two possible outcomes: success or failure. The properties of the binomial distribution are as follows:

Fixed Number of Trials (n): The number of trials n is fixed in advance.

Independent Trials: Each trial is independent of the others.

Two Possible Outcomes: Each trial results in one of two outcomes: success or failure.

Constant Probability of Success (p): The probability of success p is constant for each trial.

Discrete Outcome: The binomial random variable X represents the number of successes in n trials, which can take on only non-negative integer values.

Examples of events where the binomial distribution can be applied include:

Coin Flipping:

Each flip of a fair coin is a Bernoulli trial with two possible outcomes: heads (success) or tails (failure).

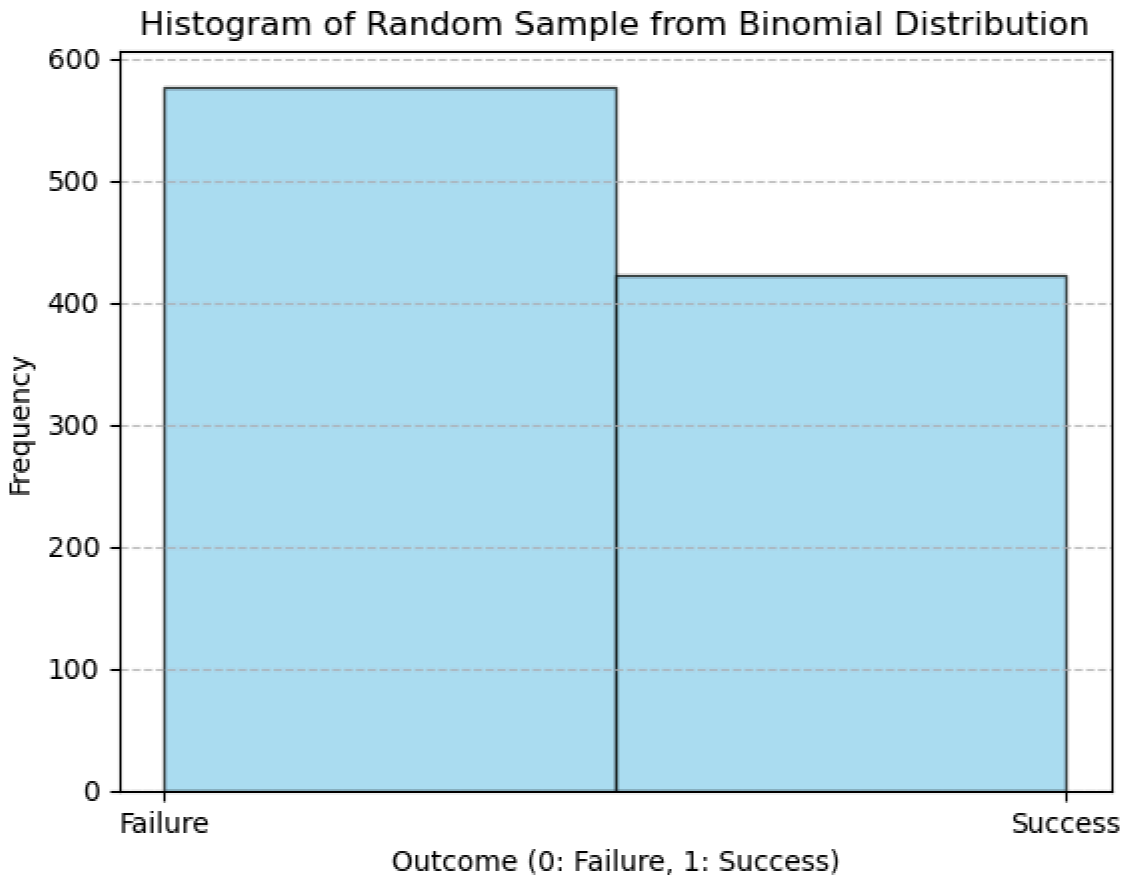
Q5. Generate a random sample of size 1000 from a binomial distribution with probability of success 0.4 and plot a histogram of the results using matplotlib.

In [2]:
```
import numpy as np
import matplotlib.pyplot as plt

# Parameters
probability_of_success = 0.4
sample_size = 1000

# Generate random sample from binomial distribution
random_sample = np.random.binomial(n=1, p=probability_of_success, size=sampl

# Plot histogram
plt.hist(random_sample, bins=2, color='skyblue', edgecolor='black', alpha=0.
plt.title('Histogram of Random Sample from Binomial Distribution')
plt.xlabel('Outcome (0: Failure, 1: Success)')
plt.ylabel('Frequency')
plt.xticks([0, 1], ['Failure', 'Success'])
plt.grid(axis='y', linestyle='--', alpha=0.7)
plt.show()
```



Q6. Write a Python function to calculate the cumulative distribution function of a Poisson distribution with given mean at a given point.

In [3]:
```
import math

def poisson_cdf(k, mean):
    """
    Calculate the cumulative distribution function (CDF) of a Poisson distri
    with the specified mean at a given point k.

    Arguments:
    k (int): The point at which to evaluate the CDF.
    mean (float): The mean of the Poisson distribution.

    Returns:
    float: The cumulative probability up to the point k.
    """
    cdf = 0
    for i in range(k + 1):
        cdf += (math.exp(-mean) * mean ** i) / math.factorial(i)
    return cdf

# Example usage:
mean = 2.5
```

```
k = 3
cdf_at_k = poisson_cdf(k, mean)
print("CDF at k =", k, ":", cdf_at_k)
```

```
CDF at k = 3 : 0.7575761331330659
```

Q7. How Binomial distribution different from Poisson distribution?

Binomial and Poisson distributions are both used to model the number of events occurring within a fixed interval or number of trials. However, they differ in several key aspects:

Nature of Events:

Binomial Distribution: It is used to model the number of successes in a fixed number of independent Bernoulli trials, where each trial has only two possible outcomes: success or failure.

Poisson Distribution: It models the number of events occurring in a fixed interval of time or space, given a constant average rate of occurrence. The events in a Poisson process are often considered rare and occur independently of each other.

Number of Trials/Interval:

Binomial Distribution: It requires a fixed number of trials (n) to be specified in advance.

Poisson Distribution: It does not require a fixed number of trials. Instead, it describes the number of events occurring in a fixed interval of time or space.

Probability of Success/Event Rate:

Binomial Distribution: It requires the probability of success (p) to be constant for each trial.

Poisson Distribution: It requires a constant average rate of occurrence (λ) for the events within the fixed interval.

Type of Events:

Binomial Distribution: It is typically used for modeling discrete events with a fixed number of trials, such as the number of heads in multiple coin flips or the number of successful sales out of a fixed number of attempts.

Poisson Distribution: It is commonly used for modeling the number of rare events occurring within a given time or space interval, such as the number of phone calls received by a call center in an hour or the number of defects in a production process

Q8. Generate a random sample of size 1000 from a Poisson distribution with mean 5 and calculate the sample mean and variance.

In [4]:
```
import numpy as np

# Parameters
mean = 5
sample_size = 1000

# Generate random sample from Poisson distribution
random_sample = np.random.poisson(mean, size=sample_size)

# Calculate sample mean and variance
sample_mean = np.mean(random_sample)
sample_variance = np.var(random_sample)

print("Sample Mean:", sample_mean)
print("Sample Variance:", sample_variance)
```

```
Sample Mean: 5.047
Sample Variance: 5.252791
```

Q9. How mean and variance are related in Binomial distribution and Poisson distribution?

Q10. In normal distribution with respect to mean position, where does the least frequent data appear?

In a normal distribution, also known as a Gaussian distribution, the least frequent data points are found in the tails of the distribution, farthest away from the mean. Specifically, the least frequent data points occur at the extreme ends of the distribution, in the regions furthest from the mean.

Since a normal distribution is symmetric, the least frequent data points are equidistant from the mean on both sides of the distribution. These points are located in the tails of the distribution, where the probability density is relatively low.

In mathematical terms, the least frequent data points are found in the tails beyond a certain number of standard deviations from the mean. This distance from the mean is determined by the standard deviation of the distribution. The further away from the mean, the lower the frequency of occurrence of data points in a normal distribution.

```
In [ ]:
```