How can the Exponential Distribution be used in Python?

Authored by

stats writer

June 28, 2024

The Exponential Distribution is a statistical distribution that is commonly used to model the time between events in a Poisson process. In Python, the Exponential Distribution can be used to generate random numbers that follow this distribution, allowing for the simulation of various real-world scenarios. This distribution can also be used for statistical inference and hypothesis testing, as well as for fitting data to the distribution and calculating probabilities. Python provides built-in functions and libraries, such as the SciPy library, that make it easy to utilize the Exponential Distribution in data analysis and modeling. By incorporating the Exponential Distribution into Python, it allows for efficient and accurate analysis of data that follows this distribution, making it a valuable tool for various applications in fields such as finance, engineering, and biology.

Use the Exponential Distribution in Python

The is a probability distribution that is used to model the time we must wait until a certain event occurs.

If a *X* follows an exponential distribution, then the cumulative distribution function of *X* can be written as:

$F(x; \lambda) = 1 - e^{-\lambda x}$

where:

λ: the rate parameter (calculated as $\lambda = 1/\mu$)e: A constant roughly equal to 2.718

This tutorial explains how to use the exponential distribution in Python.

How to Generate an Exponential Distribution

You can use the expon.rvs(scale, size) function from the SciPy library in Python to generate random values from an exponential distribution with a specific rate parameter and sample size:

```
from scipy.statsimport expon

#generate random values from exponential distribution with rate=40 and sample size=10
expon.rvs(scale=40, size=10)

array()
```

Note: You can find the complete documentation for the SciPy library .

How to Calculate Probabilities Using an Exponential Distribution

Suppose the mean number of minutes between eruptions for a certain geyser is 40 minutes. What is the probability that we'll have to wait less than 50 minutes for an eruption?

To solve this, we need to first calculate the rate parameter:

λ = 1/μλ = 1/40λ = .025

We can plug in λ = .025 and x = 50 to the formula for the CDF:

P(X ≤ x) = 1 - e-λxP(X ≤ 50) = 1 - e-.025(50)P(X ≤ 50) = 0.7135

The probability that we'll have to wait less than 50 minutes for the next eruption is 0.7135.

from scipy.statsimport expon

#calculate probability that x is less than 50 when mean rate is 40
expon.cdf(x=50, scale=40)

0.7134952031398099

The probability that we'll have to wait less than 50 minutes for the next eruption is 0.7135.

This matches the value that we calculated by hand.

How to Plot an Exponential Distribution

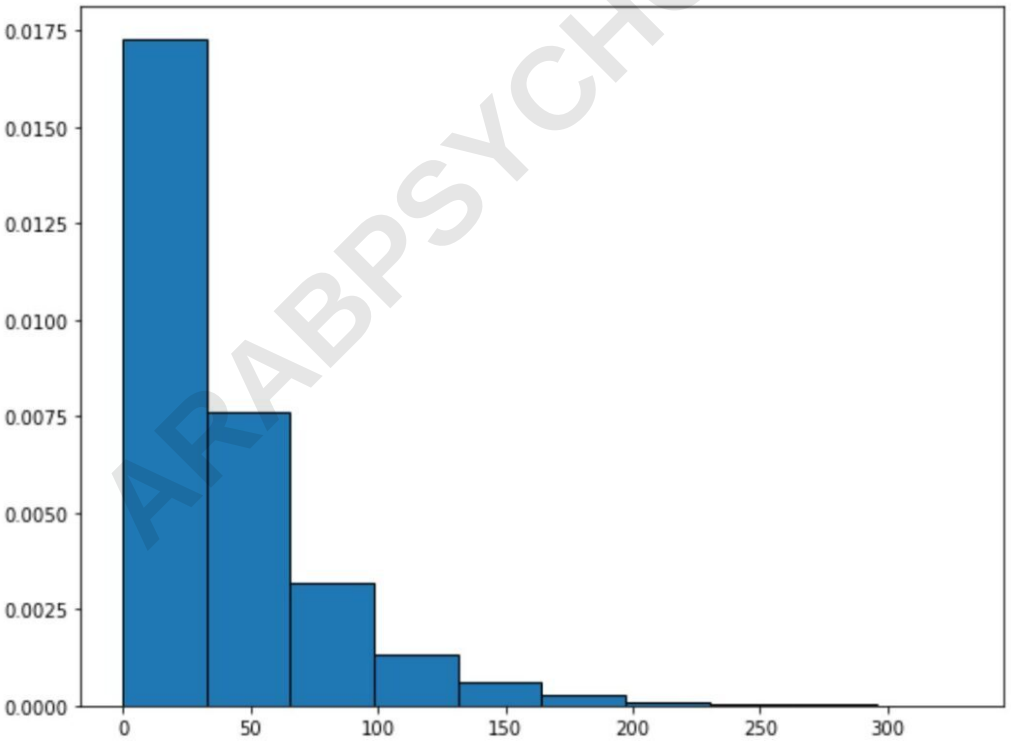
You can use the following syntax to plot an exponential

distribution with a given rate parameter:

```
from scipy.statsimport expon
import matplotlib.pyplotas plt

#generate exponential distribution with sample size 10000
x = expon.rvs(scale=40, size=10000)

#create plot of exponential distribution
plt.hist(x, density=True, edgecolor='black')
```



Additional Resources

The following tutorials explain how to use other common distributions in Python: