Thesis 5

Anna Lucia Lamacchia
Statistics

January 11, 2024

Continuous and Discrete Statistical Distributions

1 Continuous Statistical Distribution

1.1 Meaning

A continuous distribution describes the probability distribution of a continuous random variable, representing the likelihood of the variable assuming a specific value or falling within a particular interval. The key distinction lies in the fact that, in a continuous distribution, the probability associated with any single point is infinitesimally small due to the infinite number of potential values. Instead, the focus is on the probability of the variable falling within intervals, and this is captured by the probability density function (PDF).

1.2 Properties

1. **Probability Density Function (PDF)**: The core of continuous distributions is the PDF, denoted by $f(x)$. It quantifies the probability density at any given point and is integral to understanding the distribution's shape.

2. **Cumulative Distribution Function (CDF)**: The CDF, denoted by $F(x)$, represents the probability that a random variable is less than or equal to a specific value. It is the integral of the PDF and provides a cumulative view of the distribution.

3. **Support**: The range of values over which the random variable can exist is known as the support of the distribution. For continuous distributions, this support is often an unbounded interval. 1item Mean (Expectation)

and Variance: Continuous distributions have well-defined mean and variance, denoted by μ and σ^2 respectively. These measures describe the central tendency and spread of the distribution.

4. **Skewness and Kurtosis**: Skewness measures the asymmetry of the distribution, while kurtosis quantifies the shape of the tails. These properties provide insights into the distribution's departure from normality.

5. **Infinite Divisibility**: Continuous distributions possess the property of infinite divisibility, meaning that the distribution of the sum of any number of independent and identically distributed random variables from the distribution is still of the same form.

1.3 Simulation in Python

This simulation demonstrates a continuous distribution using the normal distribution. The histogram provides a visual representation of the simulated data, showcasing the characteristic shape of this distribution.

```python
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import norm

# Set the parameters for the normal distribution
mean = 0
std_dev = 1
sample_size = 1000

# Generate random samples from the normal distribution
samples = np.random.normal(mean, std_dev, sample_size)

# Create a histogram of the samples
plt.hist(samples, bins=30, density=True, alpha=0.7, label='Histogram')

# Create the theoretical probability density function (PDF)
x = np.linspace(min(samples), max(samples), 100)
pdf = norm.pdf(x, mean, std_dev)

# Plot the theoretical PDF
plt.plot(x, pdf, 'k-', label='Theoretical PDF')

# Add labels and a legend
plt.title('Simulation of Normal Distribution')
plt.xlabel('Values')
plt.ylabel('Probability Density')
plt.legend()

# Show the plot
plt.show()
```

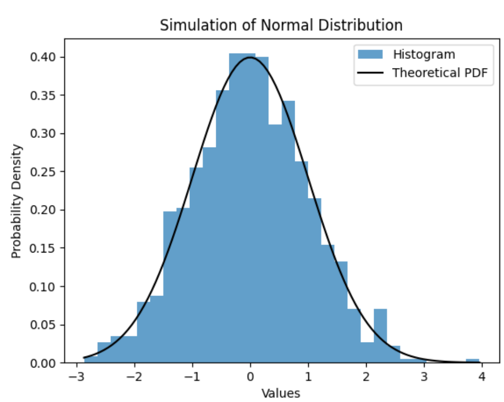

Figure 1: Simulation

2 Discrete Statistical Distribution

2.1 Meaning

A discrete distribution characterizes the probability distribution of a discrete random variable. Such variables can only assume distinct, individual values, often integers or whole numbers. The distribution is encapsulated by a probability mass function (PMF), denoted by $P(X = x)$, which provides the probability of the random variable taking a specific value x. The sum of probabilities across all possible values equals 1, reflecting the certainty that the random variable must assume one of these defined outcomes.

2.2 Properties

1. **Probability Mass Function (PMF)**: The PMF defines the probabilities associated with each possible outcome of the discrete random variable. It is a function that assigns probabilities to discrete points.

2. **Cumulative Distribution Function (CDF)**: The cumulative distribution function, denoted by $F(x)$, gives the probability that the random variable is less than or equal to a specific value x. It is obtained by summing the probabilities from the PMF.

3. **Support**: The support of a discrete distribution is the set of all possible values the random variable can take. This set is countable and often consists of integers or a specific sequence of values.

4. **Mean (Expectation) and Variance**: Discrete distributions have well-defined mean and variance, denoted by μ and σ^2 respectively. These measures describe the central tendency and spread of the distribution.

5. **Probability of Single Events**: The probability $P(X = x)$ represents the chance of a specific outcome occurring. The sum of these probabilities across all possible outcomes is 1.

6. **Skewness and Kurtosis**; While skewness and kurtosis are often associated with continuous distributions, they can still be calculated for discrete distributions to provide insights into asymmetry and tail behavior.

2.3 Simulation in Python

The script generates random samples from a binomial distribution with a given number of trials and probability of success, and then visualizes the histogram of the samples along with the theoretical probability mass function (PMF).

```python
import numpy as np
import matplotlib.pyplot as plt
from scipy.stats import binom

# Set the parameters for the binomial distribution
n_trials = 10
probability_of_success = 0.5
sample_size = 1000

# Generate random samples from the binomial distribution
samples = np.random.binomial(n_trials, probability_of_success, sample_size)

# Create a histogram of the samples
plt.hist(samples, bins=np.arange(min(samples), max(samples) + 1) - 0.5, density=True, alpha=0.7, label='Histogram')

# Create the theoretical probability mass function (PMF)
x = np.arange(min(samples), max(samples) + 1)
pmf = binom.pmf(x, n_trials, probability_of_success)

# Plot the theoretical PMF
plt.stem(x, pmf, linefmt='k-', markerfmt='ko', basefmt='k-', label='Theoretical PMF')

# Add labels and a legend
plt.title('Simulation of Binomial Distribution')
plt.xlabel('Number of Successes')
plt.ylabel('Probability')
plt.legend()

# Show the plot
plt.show()
```

Figure 2: Code

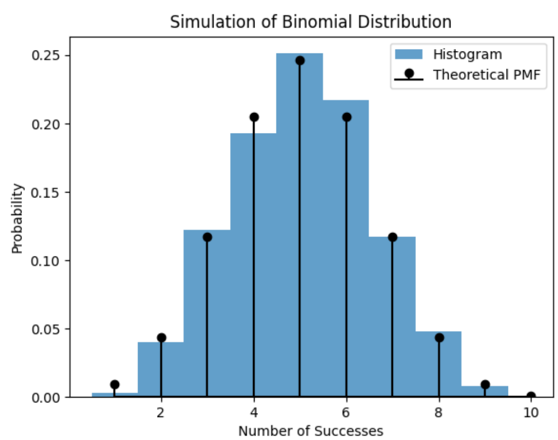

Figure 3: Chart