Contents

Chapter 1

Continuous Probability Distributions and Cumulative Density Functions

In statistics, a probability distribution describes how the values of a random variable are distributed. For continuous random variables, the probability distribution can be described by a probability density function (pdf), while the cumulative distribution function (cdf) gives the probability that the random variable is less than or equal to a certain value.

1.1 Continuous Probability Distributions

A continuous probability distribution is a probability distribution that has a pdf, which is a function that provides the probabilities of occurrence of different possible outcomes in an experiment. For a continuous distribution, the pdf $f(x)$ is such that for any two numbers a and b with $a \leq b$:

$$P(a \leq X \leq b) = \int_a^b f(x)\,dx \tag{1.1}$$

Note that for a pdf $f(x)$, we have:

$$f(x) \geq 0 \quad \text{for all } x \tag{1.2}$$

and

$$\int_{-\infty}^{\infty} f(x)\,dx = 1 \tag{1.3}$$

1.2 Cumulative Density Functions

The cumulative distribution function of a random variable X, denoted by $F(x)$, is defined as the probability that X will take a value less than or equal to x:

$$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt \tag{1.4}$$

Note that the cdf is always a non-decreasing function, and:

$$\lim_{x \to -\infty} F(x) = 0 \quad \text{and} \quad \lim_{x \to \infty} F(x) = 1 \tag{1.5}$$

1.3 Example: Normal Distribution

One common example of a continuous probability distribution is the normal (or Gaussian) distribution, characterized by its bell-shaped curve. The pdf of a normal distribution with mean μ and standard deviation σ is given by:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \tag{1.6}$$

The cdf of the normal distribution cannot be expressed in terms of elementary functions and is often computed using numerical methods. However, it can be represented as:

$$F(x) = \frac{1}{2}[1 + \mathrm{erf}(\frac{x-\mu}{\sigma\sqrt{2}})] \tag{1.7}$$

where $\mathrm{erf}(x)$ is the error function.

Chapter 2

The Normal Distribution

The Normal distribution, also known as the Gaussian distribution, is a type of continuous probability distribution for a real-valued random variable. It is one of the most important probability distributions in statistics due to its several unique properties and usefulness in many areas.

2.1 Defining the Normal Distribution

The Normal distribution is defined by its mean (μ) and standard deviation (σ). The probability density function (pdf) of a Normal distribution is given by:

$$f(x|\mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

where:

- x is the point up to which the function is integrated,

- μ is the mean or expectation of the distribution,
- σ is the standard deviation,
- σ^2 is the variance.

2.2 Importance of the Normal Distribution

There are several reasons why the Normal distribution is crucial in statistics:

- **Central Limit Theorem:** One of the main reasons for the importance of the Normal distribution is the Central Limit Theorem (CLT). The CLT states that the distribution of the sum (or average) of a large number of independent, identically distributed variables approaches a Normal distribution, regardless of the shape of the original distribution.
- **Symmetry:** The Normal distribution is symmetric, which simplifies both the theoretical analysis and the interpretation of statistical results.
- **Characterized by Two Parameters:** The Normal distribution is fully characterized by its mean and standard deviation. The mean determines the center of the distribution, and the standard deviation determines the spread or girth of the distribution.
- **Common in Nature:** Many natural phenomena follow a Normal distribution. This includes characteristics like people's heights or IQ scores, measurement errors in experiments, and many others.

Given its properties, the Normal distribution serves as a foundation for many statistical procedures and concepts, including hypothesis testing, confidence intervals, and linear regression analysis.

Chapter 3

Poisson and Exponential Probability Distributions

In this chapter, we will explore two essential probability distributions: the Poisson distribution and the exponential distribution. These distributions play a crucial role in modeling random events and phenomena, providing insights into the occurrence of events over time or in a discrete set of outcomes.

The Poisson distribution is widely used to describe the number of events that occur within a fixed interval of time or space. It is particularly useful when dealing with rare events or events that occur independently at a constant average rate. For example, the Poisson distribution can model the number of customer arrivals at a store in a given hour, the number of phone calls received by a call center in a day, or the number of defects in a production process.

The Poisson distribution is characterized by a single parameter, often denoted as λ, which represents the average rate of event occurrences in the specified interval. The probability mass function of the Poisson distribution gives the probability of observing a specific number of events within that interval.

On the other hand, the exponential distribution is used to model the time between events occurring at a constant average rate. It is commonly employed in reliability analysis, queuing theory, and survival analysis. For example, the exponential distribution can represent the time between customer arrivals at a service desk, the lifespan of electronic components, or the duration between consecutive earthquakes.

The exponential distribution is characterized by a parameter often denoted as λ, which represents the average rate of event occurrence. The probability density function of the exponential distribution describes the likelihood of observing a specific time interval between events.

In this chapter, we will explore the following key aspects of the Poisson and exponential probability distributions:

- Probability mass function and probability density function: We will dive into the mathematical representation of these distributions and learn how to calculate probabilities and densities for specific events or time intervals.
- Mean and variance: We will discuss how to calculate the mean and variance of the Poisson and exponential distributions, providing measures of central tendency and variability.
- Applications and examples: We will examine real-world scenarios where these distributions find practical applications. From analyzing customer arrival patterns to modeling equipment failure rates, we will explore a range of contexts where the Poisson and exponential distributions prove valuable.
- Relationship between the Poisson and exponential distributions: We will explore the connection between these distributions, as the exponential distribution can emerge as the waiting time between events following a Poisson process.
- Limitations and assumptions: We will also discuss the assumptions and limitations associated with the Poisson and exponential distributions, helping you understand when these models are suitable and when alternative approaches may be necessary.

By developing a solid understanding of the Poisson and exponential probability distributions, you will gain powerful tools for modeling and analyzing random events in various fields. These distributions provide valuable insights into event occurrences, time intervals, and rates, supporting decision-making processes and improving our understanding of stochastic phenomena.

Appendix A

Answers to Exercises

Index