Common Probability Distributions

Chapter: Common Probability Distributions

In this chapter, we review some well-known probability distributions in practice, both discrete and continuous, and explain some of their characteristics, such as the mean and variance.

1 Common Discrete Distributions

1.1 Uniform Distribution

A random variable X follows a discrete uniform distribution if it takes values in the finite set $E = \{x_1, x_2, ..., x_n\}$ with equal probability, i.e., $p_i = P(X = x_i) = \frac{1}{\text{card}(E)} = \frac{1}{n}$. In this case, the events $\{X = x_i\}$ are equally likely. We say that X follows a uniform distribution on E and write $X \sim \mathcal{U}_{\{x_1, x_2, ..., x_n\}}$. Thus:

- $X(\Omega) = E = \{x_1, x_2, ..., x_n\}$
- $p_i = P(X = x_i) = \frac{1}{\text{card}(E)} = \frac{1}{n}$
- $E(X) = \sum_{i=1}^{n} x_i P(X = x_i) = \sum_{i=1}^{n} x_i p_i = \frac{1}{n} \sum_{i=1}^{n} x_i$
- $E(X^2) = \sum_{i=1}^{n} x_i^2 P(X = x_i) = \sum_{i=1}^{n} x_i^2 p_i = \frac{1}{n} \sum_{i=1}^{n} x_i^2$
- $V(X) = E(X^2) - (E(X))^2$

1.1.1 Special Case

If $x_i = i$, i.e., $X(\Omega) = \{1, 2, ..., n\}$, then:

- $E(X) = \frac{1}{n} \sum_{i=1}^{n} i = \frac{n+1}{2}$
- $E(X^2) = \frac{1}{n} \sum_{i=1}^{n} i^2 = \frac{(n+1)(2n+1)}{6}$
- $V(X) = E(X^2) - (E(X))^2 = \frac{n^2-1}{12}$

1.2 Bernoulli Distribution $\mathcal{B}(p)$

Definition 1 *A Bernoulli trial is an experiment with two outcomes, often called* ***success*** *and* ***failure****.*

Consider a random experiment and an event A related to this experiment such that $P(A) = p$. Perform the experiment once, and let X be the application that takes the value 1 if A occurs and the value 0 otherwise (X is the number of occurrences of A). Then X is defined as:

$$X = \begin{cases} 1, & \text{if } A \text{ occurs} \\ 0, & \text{if } A \text{ does not occur} \end{cases}$$

The law of X is given by: $X(\Omega) = \{0, 1\}$ with $P(X = 1) = p$ and $P(X = 0) = 1-p = q$. We say that X follows a Bernoulli distribution with parameter p and write $X \sim \mathcal{B}(p)$ (or $X \sim \mathcal{B}(1, p)$).

In this case:
- $E(X) = \sum_{i=0}^{1} iP(X = i) = 0 \cdot P(X = 0) + 1 \cdot P(X = 1) = p$
- $E(X^2) = \sum_{i=0}^{1} i^2 P(X = i) = 0^2 \cdot P(X = 0) + 1^2 \cdot P(X = 1) = p$
- $V(X) = E(X^2) - (E(X))^2 = p - p^2 = p(1 - p) = pq$

1.3 Binomial Distribution $\mathcal{B}(n, p)$

Consider a Bernoulli trial that gives either A or $\overline{A}$. Perform this trial n times under the same conditions:
- The probability that A occurs is p in each trial.
- The n repetitions of the trial are independent of each other.

Let X be the number of occurrences of event A in n trials. Then X is the random variable that takes values in $\{0, 1, ..., n\}$, and its probability distribution is given by:

$$\forall k \in X(\Omega) = \{0, 1, ..., n\} : P(X = k) = C_n^k p^k (1 - p)^{n-k}$$

We say that X follows a binomial distribution with parameters n and p and write $X \sim \mathcal{B}(n, p)$.

1.4 Normal Approximation of the Binomial Distribution $\mathcal{B}(n, p)$

Let X be a random variable following a binomial distribution $\mathcal{B}(n, p)$. If n tends to infinity, the variable X tends towards a normal distribution $\mathcal{N}(m, \sigma^2)$ with $m = np$ and $\sigma^2 = np(1 - p)$. In practice, we approximate the binomial distribution with a normal distribution when n is sufficiently large ($n \geq 20$) and $np(1-p) \geq 3$.

In calculations, we use the approximation:

$$P(X = k) = P\left(k - \frac{1}{2} \leq Y \leq k + \frac{1}{2}\right),$$

where $Y \sim \mathcal{N}(np, np(1-p))$.

—

1.5 Normal Approximation of the Poisson Distribution $\mathcal{P}(\lambda)$

If $\lambda > 15$, the Poisson distribution $\mathcal{P}(\lambda)$ can be approximated by a normal distribution $\mathcal{N}(m, \sigma^2)$ where $m = \lambda$ and $\sigma^2 = \lambda$.

—

2 Common Continuous Distributions

2.1 Continuous Uniform Distribution $\mathcal{U}_{[a,b]}$

Let $a, b \in \mathbb{R}$ such that $a < b$. Let X be a real random variable with density function f. We say that X follows the uniform distribution on $[a, b]$ and write $X \sim \mathcal{U}_{[a,b]}$ if its density function f is given by:

$$f(x) = \begin{cases} \frac{1}{b-a}, & \text{if } x \in [a, b], \\ 0, & \text{if } x \notin [a, b]. \end{cases}$$

The fact that f is constant on $[a, b]$ corresponds to the idea that points are chosen in the interval $[a, b]$ with equal probability, which explains the name **uniform**.

The cumulative distribution function (CDF) of $\mathcal{U}_{[a,b]}$ is given by:

$$F(x) = \begin{cases} 0, & \text{if } x \leq a, \\ \frac{x-a}{b-a}, & \text{if } x \in [a, b], \\ 1, & \text{if } x \geq b. \end{cases}$$

Indeed:

$$\forall x \in \mathbb{R}, \quad F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt.$$

- If $x \leq a$: $F(x) = \int_{-\infty}^{x} 0\,dt = 0$, - If $a \leq x \leq b$: $F(x) = \int_{-\infty}^{a} 0\,dt + \int_{a}^{x} \frac{1}{b-a}\,dt = \frac{x-a}{b-a}$, - If $x \geq b$: $F(x) = \int_{-\infty}^{a} 0\,dt + \int_{a}^{b} \frac{1}{b-a}\,dt + \int_{b}^{x} 0\,dt = 1$.

Characteristics of $\mathcal{U}_{[a,b]}$:

$$E(X) = \frac{a+b}{2}, \quad E(X^2) = \frac{a^2+ab+b^2}{3}, \quad V(X) = \frac{(b-a)^2}{12}.$$

—

2.2 Exponential Distribution

Let $\lambda > 0$, and let X be a real random variable with density function f. We say that X follows the exponential distribution with parameter λ and write $X \sim \mathcal{E}(\lambda)$ if its density function f is given by:

$$f(x) = \begin{cases} \lambda e^{-\lambda x}, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0. \end{cases}$$

The cumulative distribution function (CDF) of $\mathcal{E}(\lambda)$ is given by:

$$F(x) = \begin{cases} 1 - e^{-\lambda x}, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0. \end{cases}$$

Indeed:

$$\forall x \in \mathbb{R}, \quad F(x) = P(X \leq x) = \int_{-\infty}^{x} f(t)\,dt.$$

- If $x \leq 0$: $F(x) = \int_{-\infty}^{x} 0\,dt = 0$, - If $x \geq 0$: $F(x) = \int_{-\infty}^{0} 0\,dt + \int_{0}^{x} \lambda e^{-\lambda t}\,dt = 1 - e^{-\lambda x}$.

Characteristics of $\mathcal{E}(\lambda)$:

$$E(X) = \frac{1}{\lambda}, \quad E(X^2) = \frac{2}{\lambda^2}, \quad V(X) = \frac{1}{\lambda^2}.$$

—

2.3 Gamma Distribution $\gamma(\alpha, \lambda)$

Let $\alpha > 0$ and $\lambda > 0$. The gamma distribution with parameters (α, λ) has the density function:

$$f(x) = \begin{cases} \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, & \text{if } x \geq 0, \\ 0, & \text{if } x < 0. \end{cases}$$

The function $\Gamma(\alpha)$ is defined for $\alpha > 0$ as:

$$\Gamma(\alpha) = \int_0^{+\infty} x^{\alpha-1} e^{-x} \, dx.$$

The parameter α is the shape parameter, and λ is the scale parameter.

Properties: 1. If T is a random variable following the gamma distribution with parameters (α, λ), then:

$$E(T) = \frac{\alpha}{\lambda}, \quad V(T) = \frac{\alpha}{\lambda^2}.$$

2. Special cases: - If $\alpha = 1$, the gamma distribution becomes the exponential distribution with parameter λ. - If $\alpha = n$ (an integer), the gamma distribution becomes the Erlang distribution of order n with parameter λ. - If $\alpha = \frac{n}{2}$ and $\lambda = \frac{1}{2}$, the gamma distribution becomes the chi-squared distribution with n degrees of freedom.

—

2.4 Cauchy Distribution $Cau(\theta, x_0)$

The Cauchy distribution, also known as the Lorentz distribution, is a classical probability distribution named after the mathematician Augustin-Louis Cauchy. A random variable X follows the Cauchy distribution if it has the density function f, depending on two parameters $x_0 \in \mathbb{R}$ (location) and $\theta > 0$ (scale), given by:

$$f(x) = \frac{1}{\pi} \frac{\theta}{\theta^2 + (x - x_0)^2} = \frac{1}{\pi\theta} \frac{1}{1 + \left(\frac{x-x_0}{\theta}\right)^2}.$$

It is clear that $\forall x \in \mathbb{R}, f(x) > 0$, and:

$$\int_{-\infty}^{+\infty} f(x) \, dx = \frac{1}{\pi} \left[\arctan\left(\frac{x - x_0}{\theta}\right) \right]_{-\infty}^{+\infty} = 1.$$

The cumulative distribution function (CDF) of $Cau(\theta, x_0)$ is given by:

$$F(x) = \frac{1}{\pi} \arctan\left(\frac{x - x_0}{\theta}\right) + \frac{1}{2}.$$

Expectation and Variance: The Cauchy distribution does not admit an expectation or variance because the functions $xf(x)$ and $x^2 f(x)$ are not integrable over $\mathbb{R}$.

—

2.5 Normal Distribution (Laplace-Gauss Distribution) $\mathcal{N}(m, \sigma^2)$

The following classical result is admitted without proof.

Theorem: Let $m \in \mathbb{R}$ and $\sigma > 0$. Then the function $f : \mathbb{R} \to \mathbb{R}$ defined by:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}$$

is a probability density function on $\mathbb{R}$.

Definition: A continuous random variable X follows the normal distribution $\mathcal{N}(m, \sigma^2)$ if its density function f is given by:

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

Properties:

$$E(X) = m, \quad V(X) = \sigma^2.$$

Standard Normal Distribution: If $m = 0$ and $\sigma = 1$, the distribution is called the standard normal distribution $\mathcal{N}(0, 1)$, with density:

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Cumulative Distribution Function (CDF): The CDF of the standard normal distribution is denoted by Φ:

$$\Phi(t) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{t} e^{-\frac{x^2}{2}}\, dx.$$

Properties of Φ:

$$\forall t \in \mathbb{R}, \quad \Phi(t) = 1 - \Phi(-t).$$

If $X \sim \mathcal{N}(m, \sigma^2)$, then the standardized variable $Z = \frac{X-m}{\sigma} \sim \mathcal{N}(0, 1)$. Thus, we can always use the standard normal table to compute probabilities.