

# Probability Distributions

## Random Variables

A random variable  $X$  is a function from the sample space to the real numbers.

$$X : S \rightarrow \mathbb{R}$$

The range of a random variable  $X$ , shown by  $\text{Range}(X)$  or  $R_X$ , is the set of possible values of  $X$ .

## Discrete Random Variables

$X$  is a discrete random variable, if its range is countable. For example, rolling a dice which results in 6 different outputs  $(1, 2, \dots, 6)$  with equally likely chances of being rolled (i.e.  $\frac{1}{6}$ ).

It has a discrete probability which is calculated on exact points.

Examples of distributions: **Binomial**, **Poisson**

**Probability Mass Function** Let  $X$  be a discrete random variable with range  $R_X = \{x_1, x_2, x_3, \dots\}$  (finite or countably infinite). The function

$$P_X(x_k) = P(X = x_k), \text{ for } k = 1, 2, 3, \dots$$

is called the probability mass function (PMF) of  $X$ .

**Cumulative Distribution Function** The cumulative distribution function (CDF) of random variable  $X$  is defined as

$$F_X(x) = P(X \leq x), \text{ for all } x \in \mathbb{R}$$

## Continuous Random Variables

$X$  is a continuous random variable, if its range is uncountable. For example, the time it takes to complete an exam is 60 minutes, possible values are all the real numbers on the interval  $[0, 60]$ .

It has a continuous probability which is measured over intervals.

Examples of distributions: **Exponential, Uniform, Normal**

**Probability Density Function** Consider a continuous random variable  $X$  with PDF  $f_X(x)$ . We have

1.  $f_X(x) \geq 0$  for all  $x \in \mathbb{R}$ .
  2.  $\int_{-\infty}^{\infty} f_X(u) du = 1$ .
  3.  $P(a < X \leq b) = F_X(b) - F_X(a) = \int_a^b f_X(u) du$ .
- 

## Expectation

If you have a collection of numbers  $a_1, a_2, \dots, a_N$ , their average is a single number that describes the whole collection. Now, consider a random variable  $X$ . We would like to define its average, or as it is called in probability, its expected value or mean. The expected value is defined as the weighted average of the values in the range.

### Definition

Let  $X$  be a discrete random variable with range  $R_X = \{x_1, x_2, x_3, \dots\}$  (finite or countably infinite). The expected value of  $X$ , denoted by  $EX$  is defined as

$$EX = \sum_{x_k \in R_X} x_k P(X = x_k) = \sum_{x_k \in R_X} x_k P_X(x_k).$$

---

## Variance

The variance is a measure of how spread out the distribution of a random variable is.

### Definition

The variance of a random variable  $X$ , with mean  $EX = \mu_X$ , is defined as

$$\text{Var}(X) = E[(X - \mu_X)^2].$$

with the computational formula

$$\text{Var}(X) = E[X^2] - [EX]^2$$

---

## Binomial Distribution

The random experiment behind the binomial distribution is as follows. Suppose that I have a coin with  $P(H) = p$ . I toss the coin  $n$  times and define  $X$  to be the total number of heads that I observe. Then  $X$  is binomial with parameter  $n$  and  $p$ , and we write  $X \sim \text{Binomial}(n, p)$ . The range of  $X$  in this case is  $R_X = \{0, 1, 2, \dots, n\}$ .

$$P_X(k) = \binom{n}{k} p^k (1-p)^{n-k}, \text{ for } k = 0, 1, 2, \dots, n$$

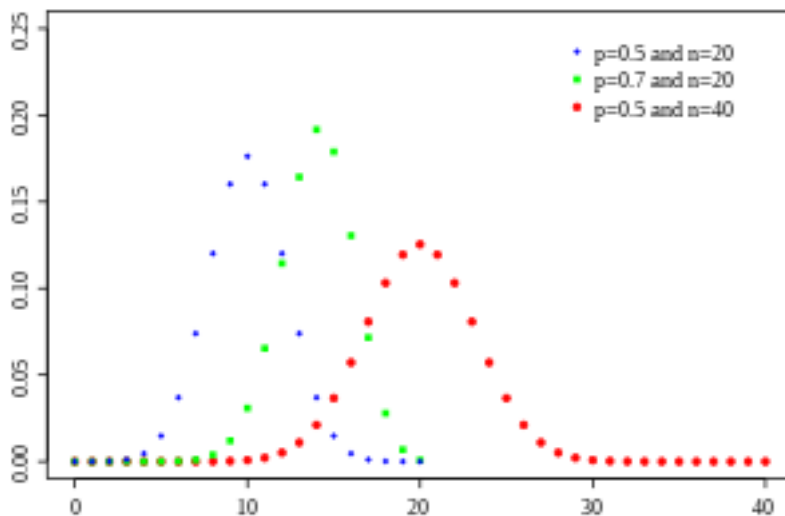


Figure 1: Binomial Distribution

### Definition

A random variable  $X$  is said to be a binomial random variable with parameters  $n$  and  $p$ , shown as  $X \sim \text{Binomial}(n, p)$ , if its PMF is given by

$$P_X(k) = \begin{cases} \binom{n}{k} p^k (1-p)^{n-k} & \text{for } k = 0, 1, 2, \dots, n \\ 0 & \text{otherwise} \end{cases}$$

where  $0 < p < 1$

### Expectation

Using the *linearity of expectation* we can deduce the expectation of a binomial distribution to be  $EX = np$ .

### Variance

$$\text{Var}(X) = np(1-p)$$

## Poisson Distribution

In practice, it is often an approximation of a real-life random variable. Here is an example of a scenario where a Poisson random variable might be used. Suppose that we are counting the number of customers who visit a certain store from 1pm to 2pm. Based on data from previous days, we know that on average  $\lambda = 15$  customers visit the store. Of course, there will be more customers some days and fewer on others. Here, we may model the random variable  $X$  showing the number customers as a Poisson random variable with parameter  $\lambda = 15$ .

### Definition

A random variable  $X$  is said to be a Poisson random variable with parameter  $\lambda$ , shown as  $X \sim \text{Poisson}(\lambda)$ , if its range is  $R_X = \{0, 1, 2, 3, \dots\}$ , and its PMF is given by

$$P_X(k) = \begin{cases} \frac{e^{-\lambda} \lambda^k}{k!} & \text{for } k \in R_X \\ 0 & \text{otherwise} \end{cases}$$

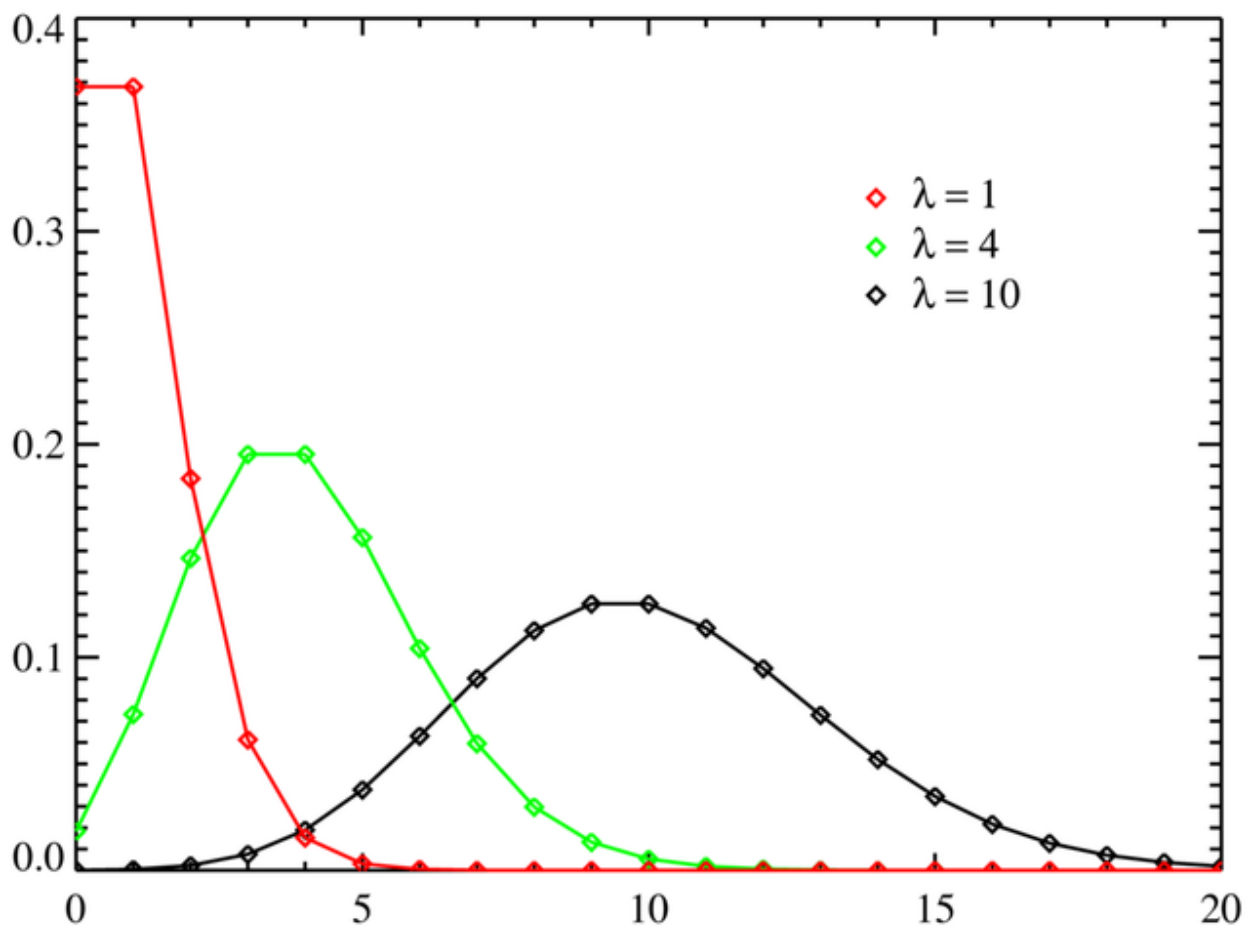


Figure 2: Poisson Distribution

**Expectation**

$$EX = \lambda$$

**Variance**

$$\text{Var}(X) = \lambda.$$

## Uniform Distribution

A continuous random variable  $X$  is said to have a Uniform distribution over the interval  $[a, b]$ , shown as  $X \sim \text{Uniform}(a, b)$ , if its PDF is given by

$$f_X(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & x < a \text{ or } x > b \end{cases}$$

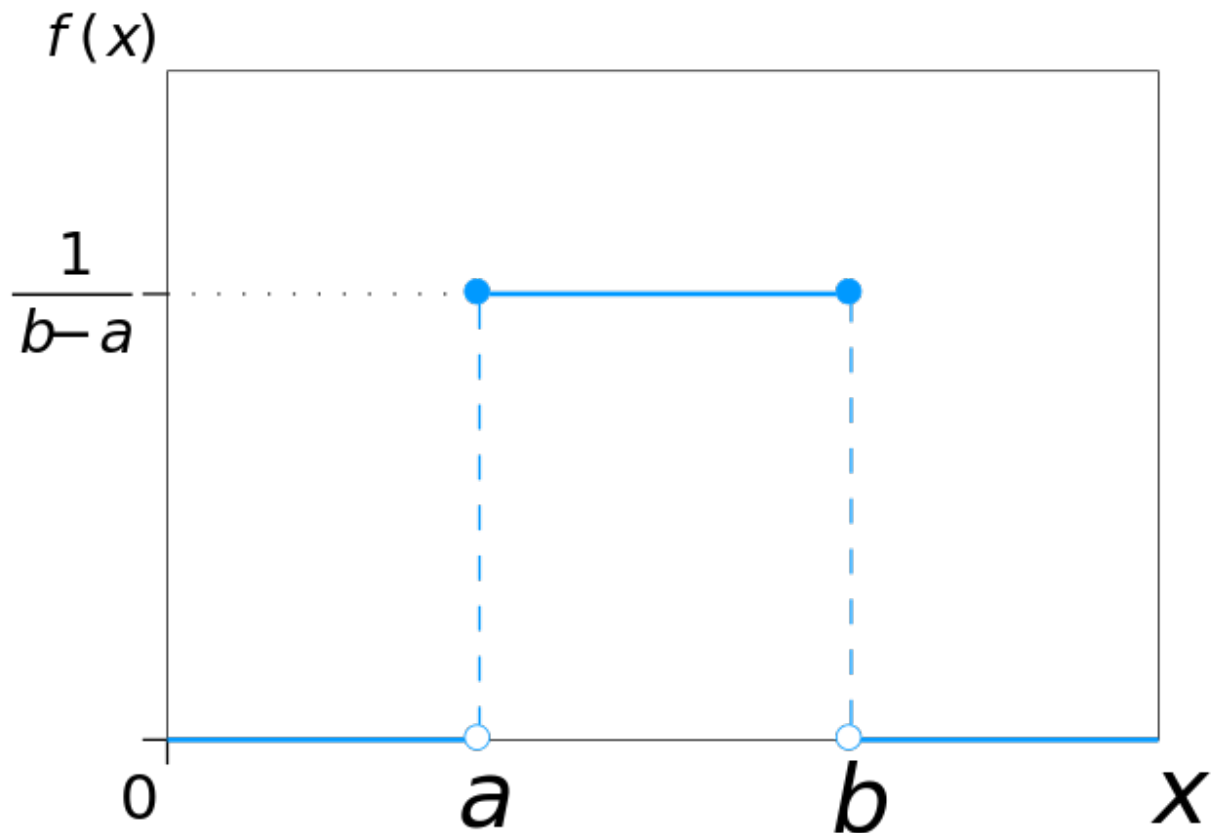


Figure 3: Uniform Distribution

**Expectation**

$$EX = \frac{a+b}{2}$$

## Variance

$$\text{Var}(X) = \frac{(b-a)^2}{12}$$

## Exponential Distribution

It is often used to model the time elapsed between events.

A continuous random variable  $X$  is said to have an exponential distribution with parameter  $\lambda > 0$ , shown as  $X \sim \text{Exponential}(\lambda)$ , if its PDF is given by

$$f_X(x) = \begin{cases} \lambda e^{-\lambda x} & x > 0 \\ 0 & \text{otherwise} \end{cases}$$

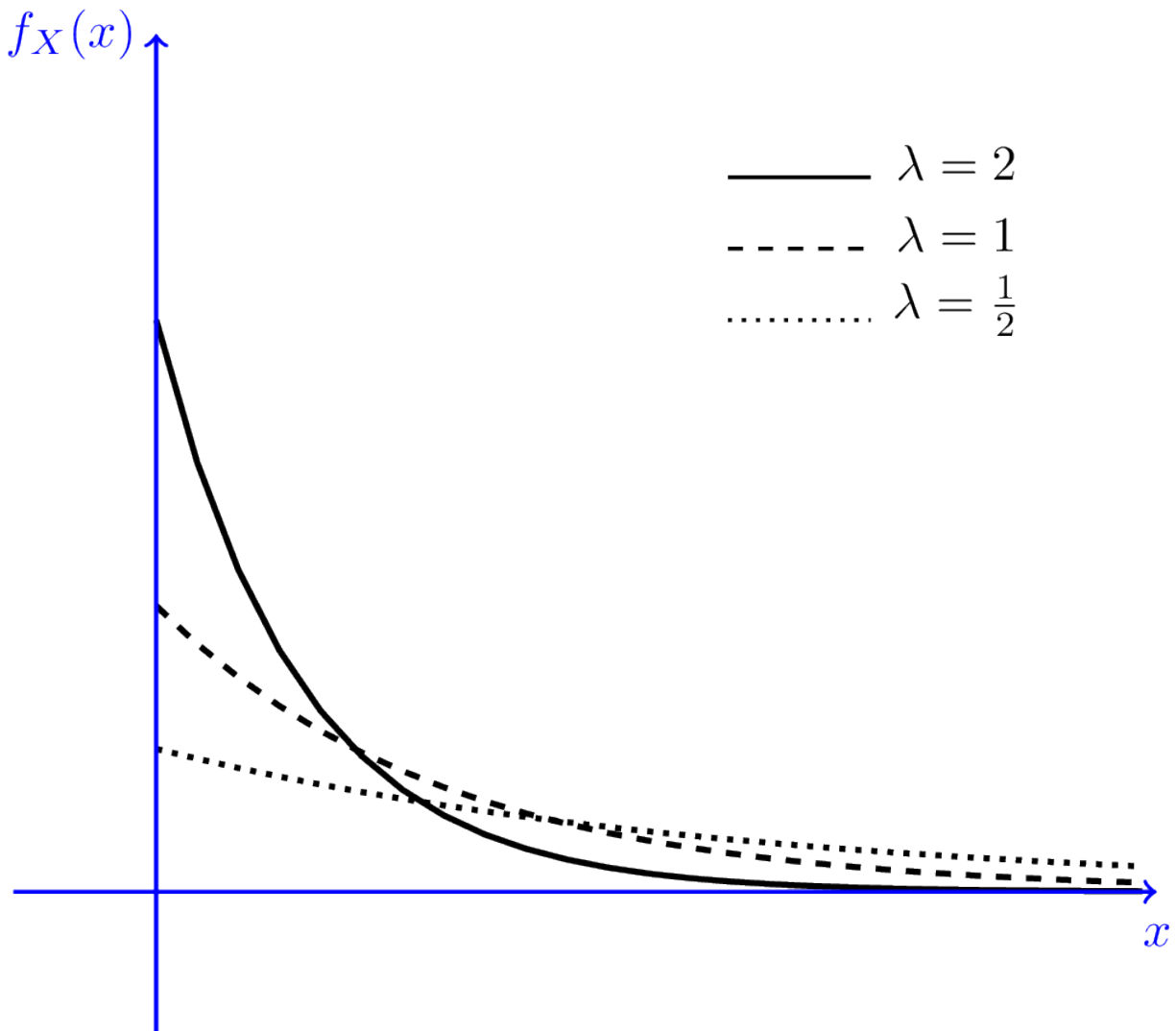


Figure 4: Exponential Distribution

## Expectation

$$EX = \frac{1}{\lambda}$$

## Variance

$$\text{Var}(X) = \frac{1}{\lambda^2}.$$

## Normal Distribution

A normal distribution, also called a Gaussian distribution, is a way of describing data that is symmetrical and bell-shaped. Imagine a bunch of measurements of people's heights. There will be more people around the average height, and fewer people who are very short or very tall. This pattern can be perfectly described by a normal distribution.

We say a random variable has a normal distribution  $X \sim N(\mu, \sigma^2)$  if its PDF is given by:

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

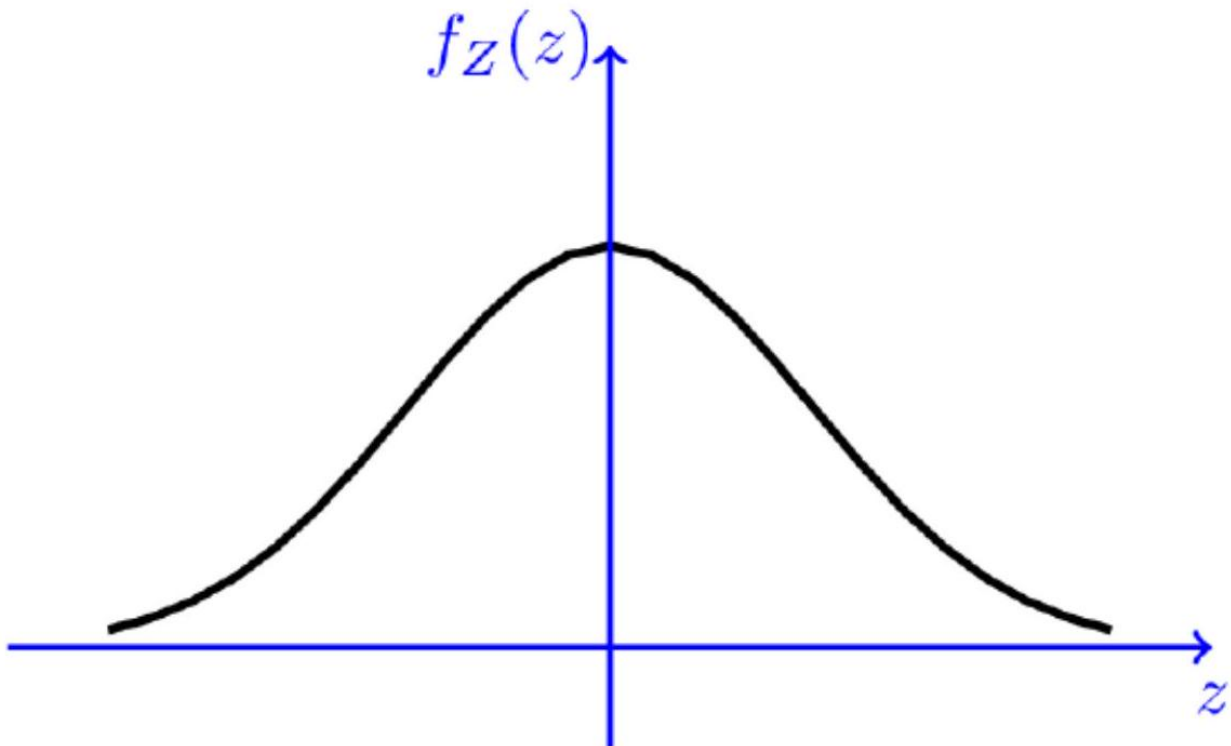


Figure 5: PDF of the standard normal random variable

**Expectation**

$$EX = 0$$

**Variance**

$$\text{Var}(X) = 1$$

—