Elements of Probability

Fall semester 2019

Keivan Mallahi-Karai
30 September 2019

Jacobs University

Recall: Expected value

If X is a discrete random variable taking values $x_1, \dots, x_n$ with probabilities $p_1, \dots, p_n$ then the expected value of X is defined by

$$\mathbb{E}[X] = p_1 x_1 + \dots + p_n x_n.$$

Compare:

- X is always equal to zero.
- Y can takes two values 1000 and -1000, each with probability $1/2$.

$$\mathbb{E}[X] = \mathbb{E}[Y] = 0.$$

How to quantify the difference in the distributions of X and Y?

Variability of a random variable

Suppose X is a random variable. How does one measure how spread-out X is?

We will do it in three steps

1. Take $X - \mathbb{E}[X]$. This moves the average to 0.
2. Squaring: consider $(X - \mathbb{E}[X)])^2$.
3. Average it: $\mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$.

Definition
The variance of a random variable X is defined by

$$\mathrm{Var}[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right].$$

Small variance means X is concentrated around its average.

Large variance means X is spread-out.

Example

Suppose X has Bernoulli distribution with parameter p. Then we have

$$\mathrm{Var}[X] = \mathbb{E}\left[(X - p)^2\right] = (-p)^2(1-p) + (1-p)^2 p = p(1-p).$$

For a binomial random variable X with parameters p and n one can show that

$$\mathrm{Var}[X] = np(1-p).$$

For a poisson random variable X with parameter λ, we have

$$\mathrm{Var}[X] = \lambda.$$

Theorem (Properties of Variance)
Let X, Y be random variables and c a constant. We have

1. $\mathrm{Var}[X] = \mathbb{E}\left[X^2\right] - \mathbb{E}[X]^2$.
2. $\mathrm{Var}[cX] = c^2 \mathrm{Var}[X]$.
3. *(Non-negativity)* $\mathrm{Var}[X] \geq 0$.
4. *(translation-invariance)* $\mathrm{Var}[X + c] = \mathrm{Var}[X]$.

Continuous random variables

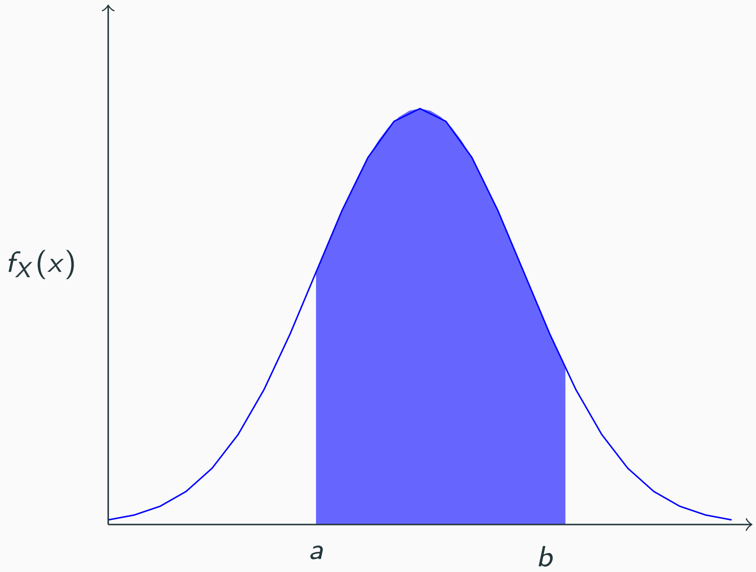
In some situations we need to deal with random quantities whose values are continuous.

Definition

A random variable $X : \Omega \to \mathbb{R}$ is called *continuous* if there exists a non-negative function $f_X : \mathbb{R} \to \mathbb{R}$, called the *probability density function* of X, such that for all values of $a < b$, we have

$$\mathbb{P}[a \leq X \leq b] = \int_a^b f_X(x) dx.$$

Integration, area, probability



The area underneath the graph of the density function from a to b represents $\mathbb{P}[a \leq X \leq b]$.

Properties of density functions

Recall that the probability mass function $p(x)$ of a discrete random variable is (a) non-negative and (b) the total probability is 1:

$$p_X(x_1) + p_X(x_2) + \cdots + p_X(x_n) = 1.$$

The probability density function of a continuous random variable satisfies the following properties:

(a) $f_X(x) \geq 0$ for all values of x.

(b) $\int_{-\infty}^{\infty} f_X(x)dx = 1.$

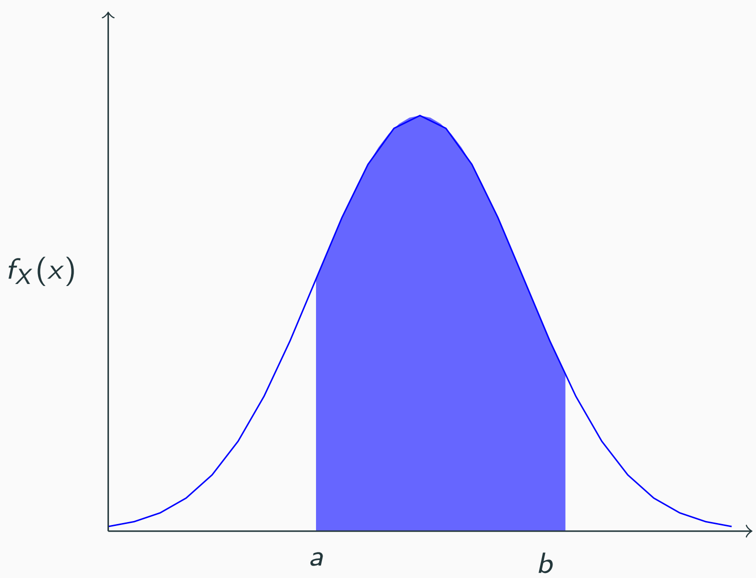
How to think about continuous random variables?

Think of choosing a random point on a line, with points on different parts of a line have different chances of being chosen.

1. Physical density vs. probability density.
2. How do we think of physical density?

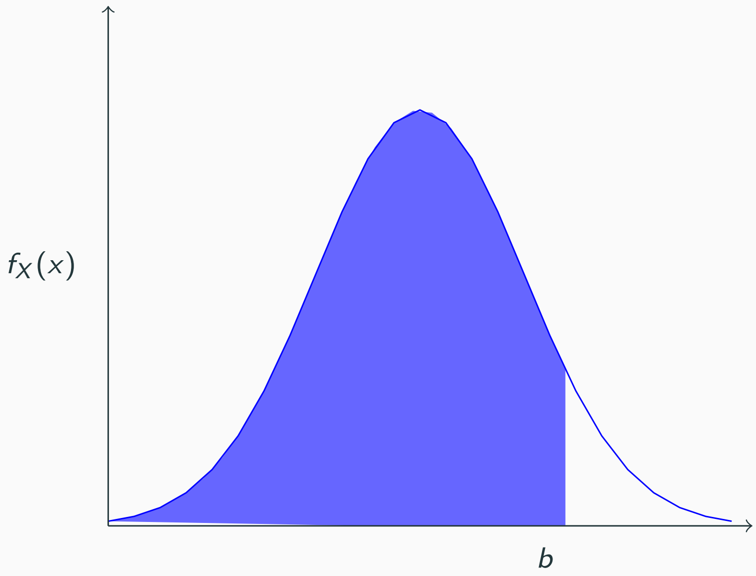
Special cases of the formula

$$\mathbb{P}[a \leq X \leq b] = \int_a^b f_X(x)\, dx.$$



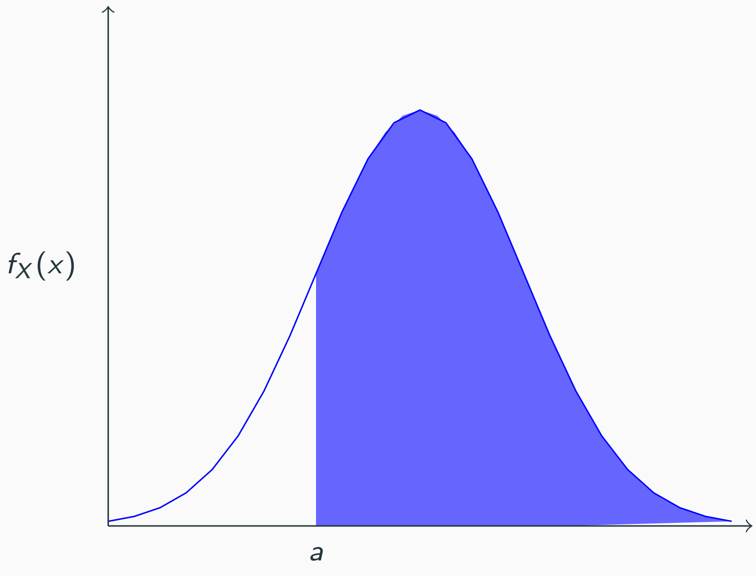
Special cases of the formula

$$\mathbb{P}[X \leq b] = \int_{-\infty}^{b} f_X(x)\, dx.$$



Special cases of the formula

$$\mathbb{P}[X \geq a] = \int_{a}^{\infty} f_X(x)\, dx.$$



The distribution function and other probabilities

Definition
The distribution function of the random variable X is given by

$$F_X(t) = \mathbb{P}\left[X \leq t\right] = \int_{-\infty}^{t} f_X(x)\ dx.$$

1.
$$\mathbb{P}\left[X \geq t\right] = 1 - F_X(t).$$

2.
$$\mathbb{P}\left[t_1 \leq X \leq t_2\right] = F_X(t_2) - F_X(t_1).$$

Relation between the probability density function and the distribution

Theorem
The relation between the probability density function and the distribution function is given by

1. *$F_X(t)$ can be obtained from the density function by integration:*

$$F_X(t) = \int_{-\infty}^{t} f_X(x)\ dx.$$

2. *$f_X(t)$ can be obtained from $F_X(t)$ by differentiation:*

$$f_X(x) = F_X^{'}(x).$$

The uniform distribution

The simplest continuous distribution is the uniform distribution.

Definition
A random variable X has uniform distribution over the interval $[a, b]$, if its probability density function is given by

$$f_X(t) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq t \leq b \\ 0 & \text{otherwise} \end{cases}$$

When we say, choose a random number between a and b, we always talk about uniform distribution.

The special case of the uniform random variable

$$f_X(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \leq x \leq b \\ 0 & \text{otherwise} \end{cases}$$

$$F_X(t) = \begin{cases} 0 & \text{if } t \leq a \\ \frac{t-a}{b-a} & \text{if } a \leq t \leq b \\ 1 & t \geq b. \end{cases}$$